Observational methods towards constraining the chemical evolution of galaxies

Alex James Cameron

ORCID: 0000-0002-0450-7306

February 2021

Doctor of Philosophy

School of Physics
Faculty of Science
The University of Melbourne

Submitted in total fulfilment for the degree of Doctor of Philosophy

Abstract

Understanding the array of physical processes that have shaped galaxy assembly remains one of the most fundamental pursuits in astrophysics. Gas in galaxies is enriched with heavy elements via stellar nucleosynthesis, but chemical abundances ("metallicity") are also shaped by galaxy-scale processes including gas accretion, feedback-driven outflows, radial gas flows, interactions, and mergers. Metallicity measurements therefore afford one of our most powerful observational probes of galaxy evolution. In this thesis I explore the performance of observational methods for constraining *(i)* gas-phase metallicity in galaxies, and *(ii)* host dark matter halo masses of galaxies; the latter of which is critical to the physics of gas flows due to its contribution to the gravitational potential well of galaxies. A particular focus is the improved understanding of systematic uncertainties near instrumental limits, which will be vital to maximise the impact of surveys conducted with future facilities.

Galaxy clustering is an efficient approach for drawing statistical connections between galaxies and their host dark matter haloes, however traditional methods are challenging to apply at $z \gtrsim 2$ where imaging survey volumes are limited. I instead apply a counts-in-cell approach to photometric $z \sim 2$ candidates from a random-pointing Hubble Space Telescope survey, showing mean counts of $N \gtrsim 5$ per field are capable of constraining the large scale galaxy bias. The James Webb Space Telescope will achieve comparable number counts out to $z \sim 8$, and thus a similar *JWST* survey could place novel constraints on the halo masses of galaxies in the epoch of reionization.

Global metallicities in low-mass galaxies afford important constraints on the impact of feedback-driven outflows on galaxy evolution. However at high-z, obtaining the requisite emission line measurements is observationally challenging. I use *Keck*/MOSFIRE spectroscopy to explore prospects for extending $z \sim 1-2$ metallicity measurements to lower masses. I find the dominant source of uncertainty arises from reduced *number* of emission lines as opposed to lower signal-to-noise, even at the detection limit. *JWST*/NIRSpec will revolutionise high-z metallicity studies due to the large suites of emission lines it will be able to assemble.

Electron temperatures (T_e) measured with auroral lines are an important baseline in metallicity studies. However the faintness of auroral lines has hitherto limited spatially resolved T_e studies. I report two separate studies based on mapping auroral lines in integral-field spectroscopy (IFS) of low-z galaxies. Measurements of auroral lines in the SAMI Galaxy Survey afford new insights into the effects of ionisation parameter variations on recovered metallicity gradients. Applying these principles to *Keck*/KCWI IFS data of an edge-on disk galaxy, I measure an extra-planar temperature gradient and present preliminary evidence for extra-planar metallicity variations.

Declaration of Authorship

I, Alex James Cameron, declare that this thesis titled, 'Observational methods towards constraining the chemical evolution of galaxies' and the work presented in it are my own. I confirm that:

- The thesis comprises only my original work towards the Doctor of Philosophy except where indicated in the preface;
- due acknowledgement has been made in the text to all other material used; and
- the thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Alex James Cameron

Date: 28/11/2020

Preface

This thesis comprises six chapters based on original work. The majority of the work presented is my own. Some aspects are the result of collaborations. Any work that is not my own is cited appropriately in the text. Chapters 2, 3, 4 & 5 are each based closely on works either already published, accepted for publication, or soon to be submitted. As first author of each of these works, I was responsible for the planning, execution and preparation of the work for publication, and contributed more than 50 % of the content of these works. Additional details are outlined as follows:

- Chapter 1 presents an original introduction and literature review of cosmology, galaxy formation and evolution, metallicity measurement techniques from emission line spectroscopy, and gas-phase metallicity studies of galaxies. This chapter is entirely my own work.

- Chapter 2 is published as:

  Cameron, Alex J.; Trenti, Michele; Livermore, Rachael C.; van der Velden, Cameron *Observational determination of the galaxy bias from cosmic variance with a random pointing survey: clustering of $z \sim 2$ galaxies from Hubble's BoRG survey*, 2019, Monthly Notices of the Royal Astronomical Society, Volume 483, Issue 2, p.1922-1933. Figure 2.1 was contributed by Rachael C. Livermore. I was primarily responsible for all other aspects of the planning, data analysis, and preparation of this work.

- Chapter 3 is published as:

  Cameron, Alex J.; Jones, Tucker; Yuan, Tiantian; Trenti, Michele; Bernard, Stephanie; Henry, Alaina; Hoag, Austin; Vulcani, Benedetta *Prospects for Extending the Mass-Metallicity Relation to Low Mass at High Redshift: A Case Study at $z \sim 1$*, 2019, The Astrophysical Journal, Volume 882, Issue 2, article id. 116. The photoionisation model grids presented in Figure 3.2 were obtained from Michael A. Dopita via private communication. I was primarily responsible for all other aspects of the planning, data analysis, and preparation of this work.

- Chapter 4 has been accepted for publication in Monthly Notices of the Royal Astronomical Society for publication as

  Cameron, Alex J.; Yuan, Tiantian; Trenti, Michele; Nicholls, David C.; Kewley, Lisa J. *Spatially resolved direct method metallicity in a high-redshift analogue local galaxy: temperature structure impact on metallicity gradients.* It was accepted on 2020 November 26. The modelling underpinning Equation 4.2 was conducted by David C. Nicholls. I was primarily responsible for all other aspects of the planning, data analysis, and preparation of this work as presented here.

- Chapter 5 is currently being prepared for publication as:

  Cameron, Alex J.; Fisher, Deanne B.; McPherson, Daniel; Yuan, Tiantian; Trenti, Michele *Mapping extra-planar electron temperature variations in MRK 1486.* This work will be submitted for publication in early 2021. The data were obtained by Deanne B. Fisher and reduced by Daniel McPherson. The photoionisation model grids presented in Figure 5.4 were obtained from Adam D. Thomas via private communication. I was primarily responsible for all other aspects of the planning, data analysis, and preparation of this work as presented here.

- Chapter 6 presents a summary of this thesis and is entirely my own work.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship, and the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. Some of this research was conducted during an overseas visit funded by a Dr Alan Kenneth Head Travelling Scholarship.

Chapter 2 is based on observations made with the NASA/ESA Hubble Space Telescope, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with programs 14701, 13767, 12905, 12572, and 11700.

The data presented in Chapters 3 & 5 were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. I also wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

The analysis presented in Chapter 4 is based on data from the SAMI Galaxy Survey Public Data Release 2 (Scott et al., 2018) which are available at https://sami-survey.org/abdr.

Much of the analysis in this thesis was carried out in Python, and variously makes use of the following Python packages: Astropy (The Astropy Collaboration et al., 2018), CMasher (van der Velden, 2020), extinction (Barbary, 2016), hoki (Stevance et al., 2020), LMFIT (Newville et al., 2014), Matplotlib (Hunter, 2007), NumPy (Harris et al., 2020), pandas (pandas development team, 2020), Photutils (Bradley et al., 2019), pPXF (Cappellari, 2017), PyNeb (Luridiana et al., 2015), and SciPy (Virtanen et al., 2020).

Acknowledgements

First and foremost, to my supervisor Michele Trenti, thank you for all of your guidance throughout the whole degree. In the first year of the degree, your open door policy was vital to helping me find my feet. In later years I spent less time in your office, but your pragmatic advice and attention to detail have always helped me to get the best out of myself and I can't thank you enough for how you've enabled me to grow as a researcher.

Secondly, to Tiantian Yuan, I've always found your enthusiasm and curiosity for research incredibly infectious and working with you has always been a genuine pleasure. So many of the ideas in this thesis grew out of long chats in the Swinburne tea room late on Tuesday afternoons and I will always fondly remember all of the lunches spent in the various cafés and eateries around Hawthorn talking about astronomy, research and whatever else came up. I wish you the best in your future endeavours and know that wherever you end up after astronomy, they'll be very lucky for having you.

I've been very lucky to be able to work with some fantastic collaborators throughout my degree. In particular, to Tucker Jones, Lisa Kewley, David Nicholls, Deanne Fisher, thank you for offering your expertise and guidance. It has undoubtedly improved the quality of the work and has really enriched my research experience. I must also thank the School of Physics, the University of Melbourne, the Laby Foundation, ASTRO 3D, and the European Southern Observatory. I've been lucky to be able to do a lot of travel throughout my degree. Thank you for supporting me to be able to have those amazing experiences.

To everyone in the Astro group, thank you for providing such a welcoming environment. Rachel Webster deserves a special mention for talking me into switching to astronomy and undertaking this degree. To Maddie, Keven, George and James, thanks for being so obliging when I needed a distraction from work. And to everyone else including Stephi, James, Julian, Nichole, Lucy, Sanjay, Hannah, Meg, Pat, Brad, Claire, Emma, Rachael and many more, thanks for many a great Friday evening spent at the Clyde.

To the gang from Cardigan Street and all the other rascals who have made life outside of uni such a blast over the last few years, you know who you are. Thanks so much for being there for me.

And finally, to Mum, Dad, Sally and Lachy, you've always been so supportive and there's no way I would have made it this far without you. A year ago I wouldn't have expected to be writing this from the farm, but with all the uncertainty that this year has thrown up, it's been so important to have had the stability and support that I've always been able to rely on from you.

Contents

List of Figures

List of Tables

Abbreviations

AGB	**A**symptotic **G**iant **B**ranch
AGN	**A**ctive **G**alactic **N**uclei
BPZ	**B**ayesian **P**hotometric Redshift (**Z**)
BoRG	**B**rightest **o**f **R**eionizing **G**alaxies survey
CEL	**C**ollisionally **E**xcited emission **L**ine
CGM	**C**ircum**g**alactic **M**edium
CLASH	**C**luster **L**ensing **A**nd **S**upernova survey with **H**ubble
CMB	**C**osmic **M**icrowave **B**ackground
DIG	**D**iffuse **I**onised **G**as
DM	**D**ark **M**atter
DRP	**D**ata **R**eduction **P**ipeline
ELT	**E**xtremely **L**arge **T**elescope
FMR	**F**undamental **M**etallicity **R**elation
FWHM	**F**ull **W**idth at **H**alf **M**aximum
GLASS	**G**rism **L**ens-**A**mplified **S**urvey from **S**pace
GOODS	**G**reat **O**bservatories **O**rigins **D**eep **S**urvey
HST	**H**ubble **S**pace **T**elescope
HUDF	**H**ubble **U**ltra-**D**eep **F**ield
ICM	**I**ntra**c**luster **M**edium
IFU	**I**ntegral **F**ield **U**nit
IGM	**I**nter**g**alactic **M**edium
IMF	**I**nitial **M**ass **F**unction
ISM	**I**nter**s**tellar **M**edium
IR	**I**nfra**r**ed
JWST	**J**ames **W**ebb **S**pace **T**elescope

KCWI	**K**eck **C**osmic **W**eb **I**mager
MOSDEF	**MOS**FIRE **D**eep **E**volution **F**ield survey
MOSFIRE	**M**ulti-**O**bject **S**pectrometer **F**or **I**nfra-**R**ed **E**xploration
MZR	**M**ass – Metallicity (**Z**) **R**elation
MLE	**M**aximum **L**ikelihood **E**stimator
NIR	**N**ear **I**nfra**r**ed
PNe	**P**lanetary **N**ebula**e**
PSF	**P**oint **S**pread **F**unction
RL	**R**ecombination **L**ine
SAMI	**S**ydney – **A**ustralian Astronomical Observatory **M**ulti-Object **I**ntegral Field Spectrograph
SDSS	**S**loan **D**igital **S**ky **S**urvey
SED	**S**pectral **E**nergy **D**istribution
SFR	**S**tar **F**ormation **R**ate
S/N	**S**ignal-to-**N**oise ratio
SNe	**S**uper**n**ova**e**
sSFR	**S**pecific **S**tar **F**ormation **R**ate
VLT	**V**ery **L**arge **T**elescope
WFC3	**W**ide **F**ield **C**amera **3**
ΛCDM	Λ – **C**old **D**ark **M**atter

Chapter 1

Introduction

The Milky Way is perhaps the most significant feature in the night sky[1]. To the naked eye it appears as a large band of diffuse light reaching from one horizon to the other. Following the advent of telescopes, astronomers were able to identify that the Milky Way consists of a disk of stars in which our Sun resides, with even early measurements suggesting its extent to be on the order of tens of kiloparsecs (Shapley, 1918). Even with the telescopes available a century ago, hundreds of so-called "spiral nebulae" could be observed (Curtis, 1918), and the suggestion arose that these nebulae were entire stellar systems, originally termed "island universes", perhaps as large as our own Milky Way galaxy (Curtis, 1917). Although the distances invoked to explain this theory seemed unfathomable at the time, we now know that these spiral nebulae are indeed separate galaxies, but also that those hundreds of galaxies catalogued a century ago were not even beginning to scratch the surface of the Universe's galaxy population.

Modern studies suggest that in fact the Milky Way is just one of as many as two trillion galaxies that reside within the observable Universe (Conselice et al., 2016). With modern telescopes, we observe that these galaxies exhibit a great deal of diversity in terms of size, colour, and morphology (e.g. Lintott et al., 2011). Already, this diversity tells us that there must be a great deal of complexity involved in the history of galaxies, giving rise to some key fundamental questions in modern astronomy: how did the first

[1] Assuming the sky is viewed from a sufficiently dark location

galaxies in the Universe form? And what physical processes have driven the assembly and evolution of galaxies from the early Universe until the present day?[2]

This thesis explores in detail observational methods focused on probing the chemical evolution of galaxies. These studies ultimately aid in building towards furthering our understanding of the latter of those two questions: what evolutionary processes have shaped the modern galaxy population? The work presented in this thesis will focus on applying observational methods near the limits of the capabilities of observational facilities. This will have important implications for surveys conducted with next generation facilities such as the James Webb Space Telescope (*JWST*) and ground-based 30 m class telescopes including the Extremely Large Telescope (*ELT*).

This chapter will introduce the theoretical basis and astrophysical context relevant to this thesis. Chapters 2, 3 & 4 contain work that is published or accepted for publication, relating to observational methods for improving our understanding of the evolution of galaxies. Chapter 5 outlines results from ongoing work being prepared for publication. A summary of the results presented in this thesis and their implications is presented in Chapter 6.

1.1 Cosmology

1.1.1 Expanding Universe

Early observations of the line-of-sight velocities of galaxies indicated that, beyond the very nearest of galaxies, all galaxies appear to be moving away from us, with recessional velocity increasing with increasing distance (Hubble, 1929). This can be described by Hubble's law,

$$v = H_0 d, \tag{1.1}$$

where v is recessional velocity, d is proper distance, and H is Hubble's constant in units of km Mpc^{-1} s^{-1}, describing the rate of this expansion. Coupled with the cosmological principle, which asserts that the Universe is homogeneous and isotropic (when viewed at sufficiently large scales), it is concluded that the Universe is expanding.

[2]Decadal plan for Australian astronomy (2016-2025): https://www.science.org.au/supporting-science/science-sector-analysis/reports-and-publications/decadal-plan-australian-astronomy-2016-25

Our current cosmological model posits that the Universe began with the "Big Bang" where the Universe existed as a singularity from which it has expanded across cosmic time. Note, that the units of Hubble's constant can be rearranged into units of inverse time. Thus, the inverse of Hubble's constant, "Hubble time" ($t_H = 1/H_0$), provides an estimate of the age of the Universe, with even early measurements of H_0 implying a Hubble time in excess of 10 Gyr. In reality, Hubble's "constant" is a misnomer since it has varied across cosmic time, and recent measurements have indicated that the expansion of space is in fact accelerating (Riess et al., 1998; Perlmutter et al., 1999). Tracing this expansion back in time, models based on modern observations predict the age of the Universe to be approximately 13.8 billion years (Planck Collaboration et al., 2016).

Defining $a(t)$ as the "scale factor" of the Universe to describe the evolution of its size with time, an alternative way of writing Hubble's constant is $H = \dot{a}/a$. The scale factor is normalised such that $0 \leq a \leq 1$ with $a = 0$ at the Big Bang, and $a_0 = 1$ representing the present value.

1.1.2 Cosmological redshift

Our main observational signal of the expansion of the Universe is the cosmological redshift of photons observed from distant objects. More precisely, we can define redshift (z) in terms of the relative change in wavelength of a photon from when it is emitted from a source (λ_{emit}) to when it is observed (λ_{obs}):

$$1 + z = \frac{\lambda_{obs}}{\lambda_{emit}}. \tag{1.2}$$

It is important to note that this is distinct from a Doppler shift which arises from relative line-of-sight velocities. While Doppler shifts do give rise to blue- and redshifts in galaxies, cosmological redshift, caused by the expansion of space, becomes increasingly dominant in more distant galaxies. The redshift of a photon emitted from a galaxy at time t and observed today can additionally be related to the scale factor as

$$z(t) = \frac{1}{a(t)} - 1. \tag{1.3}$$

With the exception of our nearest galactic neighbours, the line-of-sight redshifts observed in galaxies are dominated by cosmological redshift. As a result, in a cosmological context, redshift is often interchangeably used to refer to distance or lookback time.

1.1.3 ΛCDM model of the Universe

The "Lambda – Cold Dark Matter" (ΛCDM) model is the current consensus model in cosmology. It states that the Universe is comprised of three major components: dark energy (Λ), cold dark matter (CDM), and baryons[3].

Dark energy provides the largest contribution (∼69 %) to the energy density of the present-day Universe (Planck Collaboration et al., 2016). This "cosmological constant" represents some intrinsic property of space that exerts a negative pressure and causes the expansion of the Universe to accelerate. Baryons ("ordinary" matter) make up only ∼5 % of the energy density, while the dominant form of matter (contributing ∼26 % to the energy density) is cold dark matter. Dark matter is matter that only interacts with other matter via gravitational forces. The "cold" aspect refers to the prediction that CDM particles had relatively low velocities in the early Universe (compared to the speed of light).

1.1.4 Friedman-Lemaître-Robertson-Walker metric

We can describe the expansion of the universe with the Friedman-Lemaître-Robertson-Walker metric, which arises from solving Einstein's field equations of General Relativity (Einstein, 1916) for a homogeneous, isotropic, and expanding universe. Einstein's equations yield the following Friedman equations which describe expansion in terms of energy density:

$$\frac{\dot{a}^2 + k}{a^2} = \frac{8\pi G}{3}\rho \tag{1.4}$$

$$\frac{2\ddot{a}}{a} + \frac{\dot{a}^2 + k}{a^2} = -8\pi G P. \tag{1.5}$$

Here k describes the curvature, and ρ and P represent density and pressure respectively. High-precision curvature measurements suggest a value that is very close to flat ($k = 0$;

[3]Here this term encapsulates all "ordinary" matter, which in other contexts might be divided into baryonic and leptonic matter

Planck Collaboration et al. 2016). Accordingly we can consider the Friedman equation in a dimensionless way, relative to the density which gives us a curvature of $k = 0$, referred to as the critical density (ρ_c). Recalling that $H = \dot{a}/a$, ρ_c can be expressed as:

$$\rho_c = \frac{3H^2}{8\pi G}. \tag{1.6}$$

We can now write energy density relative to the critical density, by defining $\Omega = \rho/\rho_c$. The individual contributions of mass[4] ($\Omega_m = \rho_m/\rho_c$), radiation ($\Omega_r = \rho_r/\rho_c$) and the cosmological constant ($\Omega_\Lambda = \rho_\Lambda/\rho_c$) can be considered separately, with a total contribution:

$$\Omega = \Omega_m + \Omega_r + \Omega_\Lambda. \tag{1.7}$$

For the flat $k = 0$ Universe, we would have $\Omega = 1$. With equations of state for radiation, non-relativistic matter and a cosmological constant, we arrive at the dimensionless form of the Friedman equation below, describing the expansion of a flat universe with time t in terms of the scale factor a:

$$H(t)^2 = H_0^2\left[\Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_\Lambda\right]. \tag{1.8}$$

Using Equation 1.3, we can instead write this in terms of redshift z:

$$H(z)^2 = H_0^2\left[\Omega_m (1+z)^3 + \Omega_r (1+z)^4 + \Omega_\Lambda\right]. \tag{1.9}$$

Here H_0 refers to the present day value of Hubble's constant. Modern measurements suggest present values of $H_0 = 67.8 \pm 0.9$ km Mpc^{-1} s^{-1}, $\Omega_\Lambda = 0.692 \pm 0.012$, $\Omega_m = 0.308 \pm 0.012$ (dominated by dark matter; total baryonic contribution of $\Omega_{baryon} = 0.0484 \pm 0.0010$), and Ω_r has a negligible contribution in the present day (Planck Collaboration et al., 2016).

[4] In this definition, Ω_m comprises the contribution of both dark matter and baryonic matter.

1.2 Early cosmic history and the first galaxies

1.2.1 Early Universe

In the standard cosmological model, the beginning of the Universe 13.8 billion years ago involved a hot, dense phase during which the Universe cools as it expands rapidly from a singularity. Sub-atomic particles form within the first few seconds of cooling. After a few minutes, the Universe cools sufficiently to prevent the fusion of protons and neutrons into atomic nuclei, freezing the primordial elemental abundances to be ∼75 % hydrogen and ∼25 % helium by mass, and only trace amounts of lithium (Schramm & Turner, 1998; Olive et al., 2000; Cyburt et al., 2016).

Approximately 380,000 years after the Big Bang, the Universe cooled to a temperature of around ∼3000 K enabling protons and electrons in this primordial plasma to recombine to form neutral hydrogen atoms ("recombination"). At this point, radiation and matter in the Universe became decoupled and light could freely propagate. The radiation from this last scattering surface can thus still be observed today, albeit significantly redshifted ($z \sim 1100$) due to subsequent expansion. Present day observations of this radiation conducted in microwave wavelengths are consistent with a black-body with $T = 2.7$ K, known as the Cosmic Microwave Background (CMB) (Penzias & Wilson, 1965; Dicke et al., 1965; Planck Collaboration et al., 2016). The CMB provides a powerful probe of cosmology as it offers us the earliest view of the Universe that electromagnetic radiation can provide.

The period after recombination is often referred to as the "dark ages". During this time, baryonic matter is comprised mostly of neutral gas. At some point, the collapse of this gas under self-gravity in overdense regions leads to the formation of the first stars and galaxies, signalling the end of the dark ages. The highest redshift galaxy so far observed is at redshift $z = 11.1$ (Oesch et al., 2016), indicating that earliest galaxies must have formed by at least $t \approx 400$ Myr. Predictions from theory suggest that perhaps the first stars formed even earlier ($t \approx 30$ Myr, $z \sim 65$; Tegmark et al. 1997; Naoz et al. 2006).

Certainly, our understanding of when and how the first astronomical objects formed remains incomplete. The remainder of this section provides an overview of the current theoretical basis for the formation of these first galaxies as it is relevant to the studies of galaxy evolution presented in this thesis.

1.2.2 Origin of structure

The cosmological principle states that the Universe is homogeneous and isotropic. While this is true on the largest scales, on smaller scales it is not the case. Indeed, galaxies are one clear example of small scale inhomogeneity, having matter densities vastly exceeding that of the background universe. Spatial deviations from uniformity in CMB temperature measurements are of order 10^{-5} (Mather et al., 1990; Planck Collaboration et al., 2016), suggesting the universe was highly uniform at redshift $z \sim 1100$. However, the observation of galaxies at redshift $z \gtrsim 10$ indicates that already within ∼500 Myr, inhomogeneities have evolved to have a significant density contrast compared to the background universe.

It is postulated that cosmological structure arose from quantum fluctuations in the very early Universe (e.g. Blumenthal et al., 1984). At this time, the scale of these density fluctuations was much smaller than the average density of the Universe (that is, $\Delta\rho/\bar{\rho} \ll 1$). This was followed by a period of rapid exponential expansion ("inflation"; e.g. Linde 1983)[5], during which time the scale of these fluctuations grow rapidly, exiting the causal horizon.

The distribution of matter in the Universe is described by the power spectrum $P(k)$ which describes the power at various scales in terms of wavenumber k. Immediately following the inflationary period, the power spectrum arising from these primordial fluctuations is thought to be well approximated by:

$$P_k \propto k^n \tag{1.10}$$

with $n \approx 1$. Following the end of the inflationary era, when the expansion of space is no longer exponential, these primordial fluctuations will begin to re-enter the horizon, with the smallest fluctuations entering at earlier times. Overdensities arising from these fluctuations may begin to collapse under self-gravity, eventually giving rise to the first structures, and the power spectrum will begin to evolve from this primordial form.

[5]The inflationary model is a widely adopted extension of the ΛCDM cosmological model

1.2.3 Dark matter haloes

Recall from Section 1.1 that measurements of the background density of the Universe are consistent with a near flat Universe (curvature $k \approx 0$). If we consider a spherical region with some overdensity arising from primordial fluctuations, this region will have some density $\rho = \rho_b(1 + \delta)$, higher than the background density (ρ_b). Since for $k = 0$, the background density is the critical density (ρ_c), this higher density will instead give $k > 0$, and this spherical region is thus a closed universe. Rearranging Equation 1.4 from earlier, we have that:

$$H^2 = \frac{8\pi G\rho}{3} - \frac{k}{a^2}. \tag{1.11}$$

With positive k, the expansion of this region will be slowed relative to the background universe, eventually reaching a point in time (t_{turn}) where growth of its proper size "turns around" and it begins collapsing. Assuming the region deviates from perfect spherical symmetry[6], as dark matter particles within this region collapse they will undergo a complex array of two-body interactions, eventually reaching a steady "virialised" state. The virialised dark matter halo will obey the virial theorem which says:

$$-2E_k = E_p \tag{1.12}$$

where E_k is kinetic energy and $E_p = -0.3GM^2/R$ is the potential energy. Given that $E_{total} = E_k + E_p$, taking that $E_k = 0$ at the turn-around time t_{turn}, then from the virial theorem (1.12) we can derive that the virial radius is half of the turn-around radius ($R_{vir} = R_{turn}/2$). If we consider only dark matter, the timescale for the gravitational collapse of these over-densities to this virialised state will be of order the free-fall time ($t_{\rm ff}$). For the outermost shell of an overdense region with mass M and physical size R, this will therefore be $t_{\rm ff} = (2R^3/GM)^{1/2}$. From this we expect that smaller, more dense structures tend to collapse at earlier cosmic times. It is the widespread consensus that bound objects in the Universe form via hierarchical assembly whereby small structures form at earlier times while larger structures assemble on longer timescales via mergers of smaller objects (e.g. White & Rees, 1978).

[6]True spherical symmetry is unrealistic.

1.2.4 Dark matter halo clustering

Dark matter haloes form in over-dense regions and will thus be biased tracers of the background density field (e.g. Kaiser, 1984). This is quantified as halo bias b, with $b > 1$ indicating a propensity for haloes to preferentially sample denser regions. Dark matter haloes themselves cannot be directly observed, but galaxies that form within them will naturally also be biased tracers of the matter distribution of the Universe. Indeed, measurements of galaxy clustering have shown conclusively a preference for galaxies to reside near other galaxies (i.e. $b_g > 1$; Norberg et al. 2001; Budavári et al. 2003).

While much remains unknown about the nature of dark matter, a convenient consequence of the property that it only interacts via gravity is that the assembly of dark matter structure is relatively straightforward to predict. Indeed, many predictions have been made for the spatial distribution of dark matter haloes from the early universe to the present day, either with analytic models (e.g. Mo & White, 1996; Sheth & Tormen, 1999), or from large N-body simulations (e.g. Davis et al., 1985; Springel et al., 2005). Given these halo clustering predictions, galaxy clustering measurements afford constraints on the connection between galaxies and their host dark matter haloes (Ouchi et al., 2004, 2005; Adelberger et al., 2005; Zehavi et al., 2005; Hayashi et al., 2007). Given the relative simplicity of characterising galaxy clustering, this provides an efficient avenue toward probing the dark matter halo masses of galaxies. Later in this chapter we will see why this is critical to our understanding of the chemical evolution of galaxies.

1.2.5 Formation of the first stars and galaxies

In the broadest terms, galaxy formation occurs as the result of gravitational collapse of baryonic matter inside dark matter haloes, forming bound structures of stars, gas and dust. Unlike dark matter which can be modelled with only gravitational forces, modelling the collapse of baryonic matter must take into account a more complex array of dissipative processes.

Baryons in a collapsing cloud will be subject to an outward pressure, competing with the gravitational force that drives the collapse. As the gravitational collapse proceeds, gravitational potential energy is transformed into kinetic and thermal energy, increasing the temperature, and hence pressure, of the gas. In the absence of a mechanism for

cooling, the pressure will eventually reach an equilibrium with the gravitational force, halting further collapse and preventing star-formation from occurring (Glover, 2005).

Overcoming this barrier relies on the radiative cooling of the baryons where this thermal energy is dissipated in the form of photons. The efficiency of this radiative cooling depends on the chemical makeup of the gas (Hollenbach & McKee, 1979). The elements in this primordial gas are those arising from Big Bang Nucleosynthesis: hydrogen and helium[7]. Characteristic temperatures of gas clouds collapsing to form these first galaxies are expected to be $T \sim 100 - 1000$ K, at least an order of magnitude lower than the temperatures at which atomic hydrogen and helium can be effective coolants ($T \geq 10^4$ K) (Glover, 2005). Accordingly, molecular hydrogen (H_2) is expected to be a critical coolant during the collapse of these first objects (e.g. Mac Low & Shull, 1986), and models of primordial star formation depend on the chemistry in the primordial gas and the redshift evolution of the H_2 fraction (e.g Machacek et al., 2001).

Once the first stars are able to form, nucleosynthetic processes associated with the life-cycle of stars will create heavier elements and return them to the interstellar medium (ISM) (e.g. Heger & Woosley, 2002; Karlsson et al., 2013). This increase in chemical complexity has a fundamental impact on the formation and evolution of stars (e.g. Hollenbach & McKee, 1979; Leitherer et al., 1992), and is accordingly a critical parameter in understanding assembly of stellar mass in galaxies. Furthermore, since the chemical enrichment of the ISM is intrinsically tied to star formation and stellar evolution, chemical abundance measurements provide a powerful observational probe of how star formation, gas flows, and hierarchical assembly have shaped galaxies (e.g. Maiolino & Mannucci, 2019). The remainder of this chapter will introduce, in more detail, important concepts for studying the chemical evolution of galaxies. These will include how abundances are measured, what abundance measurements can tell us about galaxy assembly, and what key ongoing challenges remain in understanding the chemical evolution of galaxies.

[7]For the purposes of this section the trace quantities of lithium will be set aside

1.3 Chemical enrichment

1.3.1 Origin of the elements

The first atomic nuclei formed during first few minutes after the Big Bang when the Universe was very hot and dense. Big Bang Nucleosynthesis within the ΛCDM model predicts that, by mass, these primoridal elements consist of ∼75 % hydrogen, ∼25 % helium, and only trace amounts of lithium (Li/H $\sim 10^{-10}$) (e.g. Schramm & Turner, 1998; Olive et al., 2000; Cyburt et al., 2016). All heavier elements arise from nucleosynthetic processes associated with the life-cycle of stars. The astrophysical processes considered most significant to the formation of heavier elements are summarised as follows:

- **Core-collapse supernovae:** Stars with birth mass $M \gtrsim 9\ M_{\odot}$ can end their lives as a core-collapse supernova, occurring on timescales of around $t \sim 10^6 - 10^7$ yr (Janka, 2012). Alpha process fusion reactions within massive stars result in the formation of the so-called α-elements (i.e. O, Ne, Mg, Si, S, Ar, Ca). The supernova event will result in the ejection of the products back into the ISM and represents the dominant formation channel for these α-elements (Timmes et al., 1995; Kobayashi et al., 2020).

- **Asymptotic giant branch (AGB) stars:** During the last ("AGB") phase of nuclear burning in stars with initial masses $0.8 \lesssim M/M_{\odot} \lesssim 30$, a complex array of nuclear reactions and mixing occurs in the outer shells. The products of these reactions are eventually expelled as stellar winds. Significant proportions of a few light elements including carbon (∼49 %) and nitrogen (∼74 %) are produced in AGB stars (Herwig, 2005; Karakas, 2010; Kobayashi et al., 2020).

- **Type Ia supernovae:** Accretion of matter onto a CO white dwarf stellar remnant can result in explosion as a Type Ia supernova (Hillebrandt & Niemeyer, 2000). These events occur on a longer timescale after the onset of star formation ($t \sim 10^8 - 10^9$ yr; Matteucci & Recchi 2001), but are responsible for a large fraction of the so-called iron-peak elements (including Cr, Mn, Fe, and Ni). Recent modelling suggests around ∼60 % of Fe is formed via this channel (Kobayashi et al., 2020).

- **Neutron capture:** Elements heavier than iron are primarily synthesised from neutron capture processes. These are commonly divided into slow neutron capture (s-process) and rapid neutron capture (r-process), depending on the neutron density. The s-process can occur in AGB stars which provide a key formation channel for a number of heavy elements including Ba and Pb (Herwig, 2005; Kobayashi et al., 2020). There remains some debate around the site of the r-process. Neutron star mergers have been shown to provide conditions necessary for the formation of r-process elements (Wanajo et al., 2014). Despite the recent detection of a neutron star merger (Abbott et al., 2017), some suggest the rate of these mergers is too low to explain the abundances of r-process elements (e.g. Kobayashi et al., 2020) with other possible sites including magnetorotational core-collapse supernovae (Nishimura et al., 2015).

The level of heavy chemical elements at a given time in cosmic history will arise from the integrated output of each of these processes, which will be intrinsically linked to the star formation history of the Universe (Madau & Dickinson, 2014; Kobayashi et al., 2020). As a result, measurements of the abundances of elements place vital constraints on the assembly history of stellar mass in galaxies (e.g. Maiolino & Mannucci, 2019).

1.3.2 Metallicity

Astronomers use the term "metals" to encompass all elements heavier than helium, drawing a convenient distinction between primordial elements and those that are the product of stellar nucleosynthesis. Accordingly, the term "metallicity" is used to describe the overall level of chemical enrichment of heavier elements relative to primordial abundances. The strict definition of metallicity is based on the mass fraction of metals relative to hydrogen and helium. Using X, Y and Z to represent the mass fraction of hydrogen, helium, and metals respectively, this is written as:

$$Z = 1 - X - Y. \tag{1.13}$$

In practice, observational studies usually trace metallicity with specific elements, quantified as the number density of that element relative to hydrogen. In the context of

ionised gas, oxygen abundance (O/H) is widely used as a proxy for metallicity which will provide a tracer for the integrated level of chemical enrichment. While the relative abundances of different metals (e.g. O/Fe or N/O) are important in some contexts, and measurements of these can provide insight into star formation histories, for the purposes of this thesis I will focus only on tracing overall enrichment with oxygen abundances.

A key focus of this thesis is related to the methodology of making measurements of the gas-phase metallicity in galaxies, since accurate metallicity measurements are of course critical to placing constraints on the chemical evolution of galaxies. The following sections introduce how emission lines originating from ionised gas in galaxies can be used to measure metallicity.

1.4 HII regions

Gas in galaxies exists in a number of phases which can be divided into three main categories based on the state in which we observe hydrogen: molecular gas (H_2), neutral gas (atomic H; or HI), and ionised gas (H^+; or HII). While ionised gas contributes only negligibly to the gas mass of galaxies, it provides one of our most powerful probes of galaxy evolution. Much of the ionised gas in galaxies exists in what will hereafter be called "H II regions". These are the discrete nebulae[8] surrounding a recent star-formation event in which the gas is photoionised by radiation from hot, young O and B type stars (Strömgren, 1939; Osterbrock & Ferland, 2006). Since the gas in H II regions is generally thought to be representative of the gas from which these young stars formed, measurements of the metallicity of H II regions offer a snapshot in time for the galaxy, probing just this most recent generation of star-formation. One distinct advantage of studying H II regions is that the ionised gas is visible in emission and in many cases these emission lines can outshine the starlight of the galaxy. As a result, we are able to measure metallicity in H II regions out to very large cosmic distances (e.g. Maiolino et al., 2008; Sanders et al., 2020a).

[8]Note that there may be ionised gas in galaxies that it photoionised by radiation from O and B stars that is *not* associated with an H II region. An example of this might be diffuse ionised gas (DIG) or warm ionised gas that may be ionised by radiation that has 'leaked' from H II regions. Here "H II regions" refers only to the local discrete region.

1.4.1 Photoionisation of HII regions

The ionisation of hydrogen within H II regions is driven by photoionisation.

$$\mathrm{H} + \lambda \rightarrow \mathrm{H}^{+} + e^{-} \tag{1.14}$$

Here, H is a hydrogen atom in the ground state and λ is a photon with $E_\lambda \geq 13.6$ eV, corresponding to wavelengths shorter than 912 Å. The source of these ultra-violet photons are hot stars with effective temperatures of $T_{\rm eff} \gtrsim 3 \times 10^4$ K (Osterbrock & Ferland, 2006). The thermal conditions of the H II region will be determined by the equilibrium of heating from photoionisation from the stars and cooling via recombination and radiation from the nebula. Hence, the ionising flux is clearly an important parameter in terms of understanding the conditions of the nebula. This is typically described in terms of the ionisation parameter q, which is defined as the ratio of the ionising flux (Φ) to the hydrogen density (n_H), having units of cm s^{-1}. In practice, the ionisation parameter is often used in its dimensionless form U, obtained by dividing by the speed of light:

$$U \equiv \frac{q}{c} = \frac{\Phi}{n_H \cdot c} \tag{1.15}$$

Typical H II regions have ionisation parameters of $-3.5 < \log U < -2.5$ (Dopita et al., 2000), electron temperatures between 5,000 K $\leq T_e \leq$ 20,000 K, and characteristically low densities of $n_e \leq 10^4$ cm^{-3}, with $n_e \sim 10 - 100$ cm^{-3} widely observed in low-redshift nebulae (Osterbrock & Ferland, 2006). It should be noted that temperature, density and ionisation parameter are generally not constant throughout nebulae (Peimbert, 1967). Temperature and density gradients in model H II regions with idealised geometry are seen to depend on the metallicity (e.g. Figure 4 in Kewley et al. 2019b), however it bears remembering that real H II regions can have complex structures that are not easily modelled.

Metals play an important role in regulating the thermal conditions of H II regions. Certainly metals too are concomitantly photoionised by the UV flux. Furthermore, metals can be multiply ionised and may be observed in a number of ionisation states within H II regions (Kewley et al., 2019a). Metal ions including O^+, O^{2+}, N^+, S^+, S^{2+}, and

Ne^{2+} are widely observed in H II regions. Different metal ions will preferentially sample different sub-regions of the nebulae according to their ionisation potential. Higher ionisation states generally trace the inner part of the region, nearer the ionisation source, while lower ionisation states trace the outer regions (see Figure 2 in Kewley et al. 2019a).

1.4.2 Recombination lines

Emission lines provide the primary observational probe of the conditions in H II regions. The most significant emission features arise from the recombination of hydrogen nuclei with free electrons:

$$H^+ + e^- \rightarrow H + n\lambda. \tag{1.16}$$

Here $n\lambda$ represents some number of photons of precise wavelengths emitted as the electron cascades down through quantised energy levels toward the ground state in the recombined hydrogen atom. The specific wavelengths associated with these transitions give rise to the hydrogen recombination emission lines (RLs). Under Case A recombination where all emitted photons escape the nebula without re-absorption, the brightest of these will be the Lyman-α (Lyα) line at 1216 Å arising from the $n_{2\rightarrow1}$ transition. However, observable nebulae typically have large optical depths for photons arising from a transition to the $n = 1$ level (Lyman series, including Lyα). Thus, a more widely adopted assumption is Case B recombination in which all photons arising from a transition to $n = 1$ are assumed to be reabsorbed (Osterbrock & Ferland, 2006). Thus the most widely observed hydrogen recombination lines are the Balmer series lines resulting from transitions to the $n = 2$ level, which are observed at visible wavelengths. The two brightest Balmer lines are Hα and Hβ at wavelengths of 6563 Å and 4861 Å respectively. While recombination emission lines are also observed for other elements (in particular helium), since recombination emission flux is proportional to the ionic abundance, metal RLs are typically around four orders of magnitude weaker than hydrogen RLs.

1.4.3 Collisionally excited emission lines

Rather than recombination lines, a more viable observational probe of metals are the so-called "collisionally excited emission lines" (CELs). Inelastic collisions between electrons and metal ions elevate these ions into low-lying metastable excited states. Radiative transitions from these excited states are forbidden by electric dipole selection rules (but can occur by magnetic dipole or electric quadrupole transitions) meaning that the transition probabilities from these excited states are generally low and the timescale for radiative de-excitation is very long. Accordingly emission lines arising from these transitions will only be observed in gas with very low electron densities (e.g. Garstang, 1951). Historically these CELs have been referred to as "forbidden lines"; under laboratory conditions the timescale of de-excitation via emission is vastly longer than the timescale of collisional de-excitation and these lines are not observed. However, in the much lower densities of astrophysical nebulae, the timescale for collisional de-excitation is also very long. These conditions mean that, in fact, CELs are observed in abundance in the astrophysical context.

Despite the relatively low abundances of metal ions, many metal CELs are very prominent since the energies required to excited these lines are relatively small ($\sim$2 eV; Seaton 1960). Indeed the [O III] λ5007, [O III] λ4959 and [O II] $\lambda\lambda$3726, 9 CELs[9] are among the brightest emission lines observed in H II regions, often outshining hydrogen recombination lines.

It is important to note that, despite the low density of H II regions providing conditions that allow for CELs to be observed, this does not mean collisional de-excitation can be neglected. In fact, CEL fluxes remain very sensitive to the T_e and n_e of the emitting nebula. As we will see in Section 1.5, this has the important consequence of complicating the use of metal CELs to determine metallicity.

[9]A brief comment on notation: the format used here (i.e. [O III]) is common to denote nebular emission lines. The roman numeral denotes the ionisation state of the element where OI = O^0, OII = O^+, OIII = O^{2+}, and so on. Square brackets surrounding the ionic species denote it as an emission line from a collisionally excited atomic level. In this thesis, their absence indicates a recombination line.

1.4.4 Selecting HII regions: diagnostic diagrams

Before delving into how emission lines can be used to place constraints on the metallicity of H II regions, it is prudent to note that the methods introduced will all operate under an important assumption: that the primary emission line sources observed are indeed H II regions, as defined above. Emission lines can additionally originate from a number of other sources in galaxies including planetary nebulae (PNe), active galactic nuclei (AGN), shock-heated gas (shocks) and diffuse ionised gas (DIG). Since each of these has a different excitation source, each has a characteristically different spectrum, and many of the calibrations presented in this chapter do not apply to emission from these other sources.

A widely used diagnostic tool in emission line analysis is the so-called BPT diagram (Baldwin, Phillips & Terlevich, 1981). This diagram compares the emission line ratios $N2 = [\mathrm{N\,II}]\,\lambda6583/\mathrm{H}\alpha$ and $R3 = [\mathrm{O\,III}]\,\lambda5007/\mathrm{H}\beta$. Two common extensions of this diagram involve substituting the $N2$ ratio for either $S2 = [\mathrm{S\,II}]\,\lambda\lambda6716{,}31/\mathrm{H}\alpha$ or $R1 = [\mathrm{O\,I}]\,\lambda6300/\mathrm{H}\alpha$ (VO87 diagrams; Veilleux & Osterbrock, 1987). In these diagrams, H II regions are observed to inhabit a tight sequence, separated in line-ratio space from the loci inhabited by other emission sources. Various demarcation curves have been derived either from theory (e.g. Kewley et al., 2001) or from observations (e.g. Kauffmann et al., 2003) to quantify the boundary between "H II region" and "non-H II region". Observations are typically assumed to be dominated by emission from H II regions if they fall within the cut-off defined by each of these demarcation curves.

The reality of extra-galactic astronomy is that limited spatial resolution will result in blending of emission line sources. In other words, observations will rarely consist exclusively of emission from one single emission source or another. In BPT and VO87 diagrams, this presents as a 'mixing sequence' of points connecting the characteristic loci of two emission sources (e.g. H II regions and AGN), with the position along that sequence indicative of the relative contribution of the two sources. Mixing sequences are widely observed in large samples of global spectra (e.g. Kewley et al., 2006), and have also been mapped in spatially resolved observations of individual galaxies (Ho et al., 2014; Davies et al., 2014a,b, 2016, 2017; Jones et al., 2017; Zhang et al., 2017; D'Agostino et al., 2018).

Contamination from non-H II region emission sources can affect measurements of global metallicity and metallicity gradients (Sanders et al., 2017; Poetrodjojo et al., 2019). Accordingly, being able to adequately separate emission from different excitation sources and account for this contamination has important implications for chemical evolution studies. This is discussed further in Chapter 4.

1.5 Measuring abundances

Emission line fluxes depend strongly on ionic abundances. Accordingly, metallicity can be inferred using emission line flux ratios, and numerous methods exist for doing this. From a theoretical standpoint, the simplest approach is to use ratios of oxygen recombination lines to hydrogen recombination lines, as these ratios have only a weak dependence on T_e and n_e (e.g. Peimbert et al., 2017). In practice, the brightest oxygen recombination lines are prohibitively weak ($\sim 10^{-4}$ the flux of Hβ) for applications beyond the very brightest nearby astrophysical nebulae.

Instead, metallicity measurements are usually made with collisionally excited emission lines (CELs). The brightest rest-frame optical CELs for a number of metal ions, including O^+, O^{2+}, N^+ and S^+, are of comparable flux to Balmer lines. The challenge with these methods is the residual dependence of the emission line ratios on T_e and n_e which in turn are related to other H II region properties including pressure and ionisation parameter. This section introduces a number of the most widely used metallicity measurement techniques which make use of rest-frame optical collisionally excited emission lines.

1.5.1 Direct method

Since, in addition to ionic abundance, the flux of collisionally excited emission lines is sensitive to the electron temperature (T_e) and density (n_e), an appealing approach is to explicitly determine these quantities and combine those with measurements of the strong oxygen CELs to determine the metallicity. This approach is called the direct method and is widely considered a gold standard in metallicity studies.

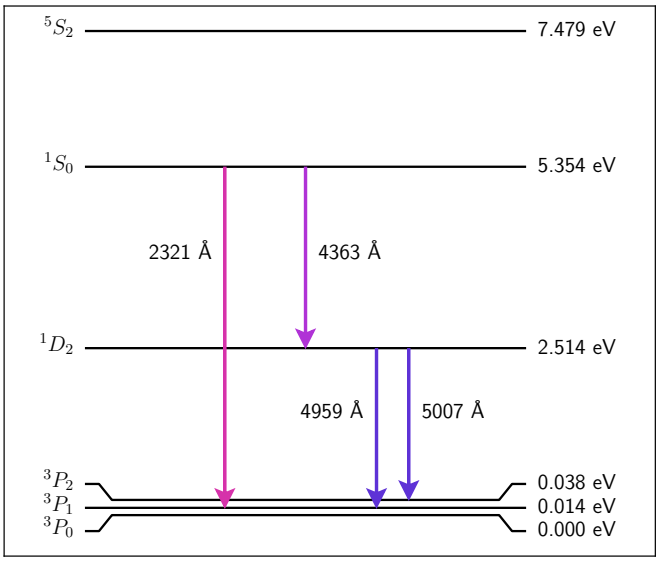

FIGURE 1.1: Lowest 6 energy levels in the O^{2+} ion. Transitions responsible for notable [O III] emission lines are indicated.

Figure 1.1 depicts the lower energy levels of the O^{2+} ion, indicating the transitions that give rise to prominent [O III] lines. By far the brightest of these is the doublet arising from transition from the $^1D_2$ to the $^3P_2$ and $^3P_1$ levels, corresponding to the [O III] lines at 5007 Å and 4959 Å respectively. Transition from the $^1S_0$ to the $^1D_2$ state results in a so-called auroral line at 4363 Å which is typically of order 10^{-2} times the flux of the brighter [O III] lines (Nicholls et al., 2020). Since this auroral line originates from a transition from a higher energy level, the specific flux ratio of this auroral line to the strong lines will be sensitive almost exclusively to the temperature[10]. Therefore, this line ratio can be used as an electron temperature diagnostic.

If we assume the low-density limit where collisional de-excitation can be ignored, then any one excitation to the $^1D_2$ level will result in emission of a λ5007 or λ4959 photon. Similarly an excitation to the $^1S_0$ level results either in one λ2321 photon, or one λ4363 photon followed by another photon with either λ5007 or λ4959. Intuitively, we can imagine that the flux ratio between the λ4363 line and the λ4959 and λ5007 lines depends on the relative energies required to excite electrons into these levels, the relative probabilities of transition (radiative de-excitation), the energy gap between these energy levels, and of course the electron temperature. Following Osterbrock & Ferland (2006),

[10]This is true for densities well below the critical density where the rate of radiative de-excitation is equal to the rate of collisional de-excitation. The critical density for [O III] λ5007 is $n_e = 7.0 \times 10^5$ cm^{-3} (Nicholls et al., 2020), which is much higher than the typical densities observed in H II regions

we can express this ratio of in terms of those parameters as:

$$\frac{j_{\lambda 4959} + j_{\lambda 5007}}{j_{\lambda 4363}} = \frac{\Upsilon_{(^3P,^1D)}}{\Upsilon_{(^3P,^1S)}} \times \frac{A_{(^1D_2,^3P_2)} \cdot \nu(\lambda 5007) + A_{(^1D_2,^3P_1)} \cdot \nu(\lambda 4959)}{\left[A_{(^1D_2,^3P_2)} + A_{(^1D_2,^3P_1)}\right] \cdot \nu(\lambda 4363)} \times \frac{A_{(^1S,^1D)} + A_{(^1S,^3P)}}{A_{(^1S,^1D)}} \times \exp\left(\frac{\Delta E}{kT_e}\right) \quad (1.17)$$

where $j_{\lambda x}$ is the flux of an emission line with wavelength x Å, with corresponding wave number $\nu(\lambda x)$, $\Upsilon_{(X,Y)}$ is the thermally averaged collision strength for excitation from $X \rightarrow Y$, $A_{(Y,X)}$ is the transition probability for $Y \rightarrow X$, ΔE is the energy gap between the $^1D_2$ and $^1S_0$ levels, and T_e is the electron temperature.

While this is a good approximation across the typical range of densities observed in H II regions, collisional de-excitation becomes more important more quickly for the stronger lines. Accounting for this density dependence and substituting numerical values in, Osterbrock & Ferland (2006) give:

$$\frac{j_{\lambda 4959} + j_{\lambda 5007}}{j_{\lambda 4363}} = \frac{7.90 \exp(3.29 \times 10^4 / T_e)}{1 + 4.5 \times 10^{-4} \cdot n_e / T_e^{1/2}}. \quad (1.18)$$

1.5.1.1 Direct method: determining T_e

Equation 1.18 describes how the ratio of [O III] $\lambda\lambda$4959, 5007 to [O III] λ4363 changes with temperature and density. This is essentially the basis of the direct method: the ratio of an auroral line to the corresponding strong-line(s) can be used to determine the electron temperature. In addition to O^{2+}, similar auroral line ratios exist for a range of ions including O^+ ([O II] $\lambda\lambda$3726, 9 / $\lambda\lambda$7320, 30), N^+ ([N II] $\lambda\lambda$6548, 83 / λ5755), S^+ ([S II] $\lambda\lambda$6716, 31 / $\lambda\lambda$4069, 76), and S^{2+} ([S III] $\lambda\lambda$9069, 9532 / λ6312).

We see in Equation 1.18 a dependence on density. Density can be well constrained by the intensity ratio of two lines from the same ionic species that have very similar excitation energy, such that the flux ratio will depend mostly on the collisional de-excitation rate, and hence the electron density (refer to Chapter 5.6 in Osterbrock & Ferland 2006). The most widely used examples are the [O II] λ3729 / λ3726 ratio and the [S II] λ6716 / λ6731 ratios (e.g. Sanders et al., 2016).

Implementations of the direct method usually involve using an iterative approach to solve simultaneously for T_e and n_e, given observed values for a temperature- and a

density-sensitive line ratio. Examples include Izotov et al. (2006) (Eq. 1 & 2), the TEMDEN routine from IRAF (Shaw & Dufour, 1995), and the `getCrossTemDen` routine from `PyNeb` (Luridiana et al., 2015).

One issue with this approach is that the ionic species most commonly used to probe temperature ([O III]) and density ([S II]), generally do not originate from the same region of the nebula (Kewley et al., 2019a). Recently, Nicholls et al. (2020) argued that since temperature and density are expected to vary within H II regions, any gains made by explicitly accounting for the relatively weak n_e dependence of the [O III] λ4363 / $\lambda\lambda$4959, 5007 ratio are likely offset by the modelling uncertainty surrounding density fluctuations within nebulae. Nicholls et al. (2020) instead propose that within reasonable expectations of the modelling uncertainty of H II regions, the density dependence can safely be ignored (provided $n_e \leq 10^3$ cm^{-3}). Accordingly, they provide a simplified method calibrated to modelling data, where T_e can be obtained from the following rational polynomial function:

$$\log_{10}(T_e) = \frac{3.5363 + 7.2939x}{1.0 + 1.6298x - 0.1221x^2 - 0.0074x^3} \tag{1.19}$$

where $x = \log_{10}(j_{\lambda4363}/[j_{\lambda4959} + j_{\lambda5007}])$, and T_e is in units of K.

Finally, it should be considered that in reality, the temperatures we measure represent something of a flux-weighted average temperature of the observed region. The spectra we obtain in extra-galactic observations will often lack the spatial resolution to resolve individual H II regions from each other, let alone be able to resolve small-scale temperature fluctuations within H II regions. Thus, we must consider that there is likely little to be gained from employing overly complicated methods to compensate for this minor density dependence.

1.5.1.2 Direct method: determining metallicity

With measurements of T_e and n_e, ion abundances can be determined from the ratios of the strong CELs with Hβ ([O III] $\lambda\lambda$4959, 5007 / Hβ and [O II] $\lambda\lambda$3726, 9 / Hβ in the cases of O^{2+} and O^+ respectively). A number of analytic relations have been derived for determining ion abundances from T_e measurements and line flux ratios (e.g. Izotov et al., 2006; Pérez-Montero, 2017). Alternatively abundances can be determined based

on atomic data with packages such as `PyNeb` (Luridiana et al., 2015). Any differences between these different methods are likely smaller than the systematic uncertainties.

Total oxygen abundance is usually derived under the assumption that the contribution of neutral oxygen (O^0) and oxygen in ionisation states of O^{3+} or higher are negligible and hence $O/H \approx O^+/H^+ + O^{2+}/H^+$ (e.g. Berg et al., 2020). Since temperature is not constant within H II regions, O^+ and O^{2+} should each be derived using the temperature of the ionisation zone from where those oxygen lines are originating (T_e[O II] and T_e[O III], respectively). Ideally, one would derive each of these temperatures independently from measurements of the [O II] and [O III] auroral lines respectively. However, the faintness of the required auroral lines makes this observationally challenging. Thus, a widely used alternative is to infer T_e([O II]) from T_e([O III]) via some assumed relation ("semi-direct" method; Campbell et al., 1986; Izotov et al., 2006; López-Sánchez et al., 2012; Andrews & Martini, 2013; Pérez-Montero, 2017; Yates et al., 2020). However, there remain unquantified uncertainties in this approach since the temperature structure can deviate from expected relations (e.g. Yates et al., 2020), and the preferred method remains the use of auroral lines from multiple species, rather than semi-direct methods. Work presented in Chapter 4 explores this issue in much more detail, and further discussion is deferred to that chapter.

Limitations of the direct method: The primary issue with the direct method is its requirement for the faint auroral lines. Even within the local Universe, detections of auroral lines in individual galaxies are largely limited to low metallicity targets, with high metallicity extensions only possible with stacking (Andrews & Martini, 2013; Curti et al., 2020a). Modest samples have been assembled at intermediate redshifts ($z \sim 0.7$; Jones et al., 2015b; Ly et al., 2016a,b; Calabrò et al., 2017) while only a handful of auroral line detections exist beyond $z \gtrsim 1$ (Patrício et al., 2018; Sanders et al., 2020b). Beyond measuring rest-optical auroral lines, Jones et al. (2020) outlined a T_e-based method instead utilising far-IR oxygen lines ([O III] 52μm and [O III] 88μm). While this is expected to be favourable beyond $z \gtrsim 5$ where these lines can be observed at millimeter wavelengths, it is difficult to apply at low-redshift. Ultimately, in order to probe the chemical evolution of galaxies, we need to be able to place constraints on the metallicity without explicitly determining T_e.

1.5.2 Strong-line methods

Strong-line methods provide an alternative approach in which ratios of the brightest emission lines are used to determine the metallicity empirically. These have the advantage of a vastly expanded range of application as compared to the direct method, however are more susceptible to variations in other physical conditions in the observed nebulae (Kewley et al., 2019a; Maiolino & Mannucci, 2019).

The simplest form of a strong-line method is to construct a strong-line diagnostic. Generally speaking, these allow for metallicity to be determined as $\log(O/H) = F(R)$ where R is a metallicity sensitive line ratio and F is some simple linear or polynomial relation. Widely used emission line ratios include (R_{23}; [O II] +[O III])/Hβ), ($O3N2$; ([N II]/Hα)/([O III]/Hβ)), and ($N2O2$; [N II]/[O II]) (Maiolino & Mannucci, 2019).

Due to the observational limitations of the direct method, the development of photoionisation modelling has been critical to our understanding of the behaviour of emission lines, particularly in high-metallicity systems. In these models, atomic data for emission line collision strengths (e.g. Lennon & Burke, 1994; Palay et al., 2012; Storey et al., 2014) is used to predict emission line fluxes over a grid of different nebular properties, given an input ionising stellar population spectrum. Typically the nebular properties of interest are metallicity and ionisation parameter, but can also include other parameters such as pressure. Two widely-used photoionisation modelling codes are available: Cloudy (Ferland et al., 2017) and Mappings (Sutherland & Dopita, 2017; Kewley et al., 2019b). A more detailed overview of photoionisation modelling can be found in the recent review by Kewley et al. (2019a).

Given the limitations of the direct method, metallicity diagnostic calibrations are typically supplemented with photoionisation model predictions either in a hybrid manner (McCall et al., 1985; Pettini & Pagel, 2004; Maiolino et al., 2008) or by moving to a fully theoretical calibration (McGaugh, 1991; Kewley & Dopita, 2002; Kewley et al., 2019a). More recently, stacking spectra of high metallicity galaxies has enabled calibrations to be derived based only on direct method measurements (Curti et al., 2017, 2020a). Numerous combinations of strong-line ratios have been calibrated in these ways, resulting in a wide variety of different strong-line diagnostics available to the modern user.[11]

[11] A snapshot of these can be found in Table 1 of Maiolino & Mannucci 2019

1.5.2.1 Challenges with strong-line methods

A critical challenge in chemical evolution studies is the long known disagreement between different metallicity measurement techniques (Kewley & Ellison, 2008; Moustakas et al., 2010; Morales-Luis et al., 2014; Maiolino & Mannucci, 2019). For example, Kewley & Ellison (2008) observed that the normalisation of the Mass-Metallicity Relation can change by more than 0.5 dex just by using different metallicity indicators.

This is due, at least in part, to the fact that none of these strong-line ratios is sensitive only to the metallicity. Indeed, observed ratios of these strong emission lines will be sensitive to variations in a range of other nebular parameters including: ionisation parameter, ISM pressure, and relative abundance ratios of different metals (e.g. N/O). This list is by no means exhaustive and all strong-line diagnostics are affected by these factors in some way or another.

Photoionsation modelling enables us to understand these dependencies to an extent. In particular, since ionisation parameter is one of the most important parameters in terms of the emission spectrum of H II regions, some recent methods mitigate ionisation parameter dependence by solving iteratively for both metallicity and ionisation parameter (e.g. Kewley et al., 2019a).

Taking this a step further, one can in principle fit arbitrary sets of emission lines to model grid predictions, simultaneously constraining metallicity alongside other nebular properties including ionisation parameter. Several publicly available implementations of this include: `IZI` (Blanc et al. 2015; O/H and U), `NebulaBayes` (Thomas et al. 2018; O/H, U and ISM pressure) and HII-CHI-MISTRY (Pérez-Montero 2014; O/H, U and N/O).[12] With the continued development of photoionisation models, interest in this sort of approach is growing. However, our ability to produce model spectra that match observations is still improving and at present these remain a complementary technique.

1.5.2.2 Strong-line methods at high-redshift

The ultimate goal of chemical evolution studies is to understand how the metallicity of galaxies has evolved across cosmic time which in turn informs our understanding of

[12]Parameters in parentheses indicate which among metallicity (O/H), ionisation parameter (U), ISM pressure, and nitrogen-to-oxygen abundance ratio (N/O) each package computes.

galaxy assembly. Naturally, an important aspect of this pursuit is to be able to make accurate measurements of the metallicity of high-redshift galaxies. Inevitably, high-redshift metallicity studies will be the domain of strong-line methods as auroral lines are simply too faint to be widely applied beyond the local Universe.

High-redshift studies are particularly susceptible to uncertainties caused by variations in ISM conditions. The reason for this is that, beyond "random" variations in ISM conditions that we may observe within a particular population of galaxies, high-redshift galaxies may have *systematically* different ISM conditions to local galaxies due to evolution across cosmic time (Kewley et al., 2013; Steidel et al., 2014; Sanders et al., 2016, 2020b; Strom et al., 2017, 2018; Topping et al., 2020a,b). Indeed ionisation parameter and electron density are observed to evolve across cosmic time (e.g Sanders et al., 2016). Additionally, the locus of high-redshift star-forming galaxies on the BPT diagram is observed to be offset from those seen in local galaxies (Kewley et al., 2013; Steidel et al., 2014). Some authors have found that this is can explained by evolution in O/Fe abundance ratios with redshift, suggesting that the shape of the ionising spectra that power H II regions will have evolved with redshift (Strom et al., 2017; Topping et al., 2020a,b). This has prompted some authors to advocate for the use of redshift-variable strong-line method calibrations to account for these effects (Bian et al., 2018; Sanders et al., 2020b,a). Certainly, it is paramount that metallicity determinations at high-redshift are robust against possible redshift evolution in the nebular conditions of galaxies if a self-consistent picture of chemical evolution is to emerge. Larger samples of auroral line measurements at high-redshift enabled by future facilities such as *JWST*/NIRSpec will no doubt help to provide some improved clarity. Further discussion around this topic is deferred to Chapter 3.

1.6 Galaxy Assembly

Studies of the cosmic star formation history show that the star formation rate density of the Universe steadily increased from the early epoch of galaxy formation for the first ~ 3.5 Gyr of cosmic time, peaking at around $z \sim 2$, before declining exponentially at $z < 1$ (e.g. review by Madau & Dickinson 2014). The sites of this star formation are, of course galaxies, and measurements of the stellar mass function of galaxies indicate that the cosmic density of galaxies at fixed stellar mass has steadily increased since at least

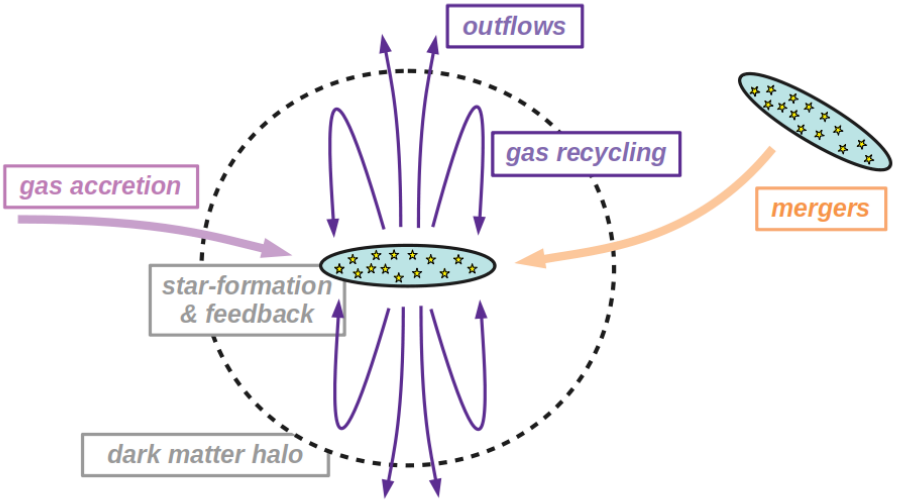

FIGURE 1.2: Simple schematic showing some key processes that can shape the evolution of galaxies. The central galaxy here depicts a disk viewed edge-on.

$z < 4$ for galaxies with $\log(M_*/M_\odot) > 10$ (Marchesini et al., 2009). In parallel with the assembly of stellar mass, since star formation is associated with the production of metals, the metal content of the Universe is also expected to increase monotonically with cosmic time (Madau & Dickinson, 2014). However, informative as these cosmological studies are about the global assembly of stellar mass, they do little to illuminate which processes shape the assembly of mass and structure within individual galaxies.

The formation of stars in galaxies will inject energy back into the surrounding ISM, introducing "feedback" which may suppress further star formation. Thus, from even a brief consideration of second order effects, it becomes clear that the assembly of stellar mass in galaxies is not so simple as the continuous collapse of cold gas in dark matter haloes forming progressively more stars. Rather, the evolutionary history of a galaxy is mediated by the interplay of a number of secular and environmental processes, which are introduced in this section.

1.6.1 The galactic fountain

Figure 1.2 presents a simple schematic depiction of the different processes that can influence the growth of a galaxy. As outlined in Sections 1.2.3–1.2.5, the first generation of stars in our model galaxy form from the collapse of primordial gas inside a virialised

dark matter halo. There is some thought that the initial burst of star formation from primordial gas was brief, with second generation star formation from enriched gas quickly becoming important (e.g. Karlsson et al., 2013).

Once some level of metal enrichment has occurred, star-formation events will give rise to stars over the approximate mass range $0.1 \lesssim M/M_{\odot} \lesssim 100$, the distribution of which is given by an initial mass function (IMF) (e.g. Salpeter, 1955; Chabrier, 2003). As discussed in Section 1.3, massive stars will further enrich the surrounding ISM with metals either by stellar winds or supernovae events. However, in addition to metals, these events also inject energy and momentum into the surrounding ISM. This provides "feedback" against further star formation by preventing the collapse of cold gas. This stellar feedback is important in galaxy models reproducing observed cosmic densities of faint galaxies (e.g. Benson et al., 2003). Beyond stellar feedback, other forms of feedback can be important in regulating star formation, including that from active galactic nuclei (AGN) in massive haloes (e.g. Croton et al., 2006).

Beyond simply suppressing further star formation via gas heating, this stellar feedback is capable of creating "galactic winds" in which gas is liberated from the disk of the galaxy (Heckman et al., 2000; Veilleux et al., 2005; Hopkins et al., 2012; Rupke, 2018). Indeed, galactic winds are ubiquitous in low-redshift galaxies with high star formation rate surface densities (Heckman, 2002), and among high-redshift galaxies (e.g. Weiner et al., 2009) when the cosmic SFR was higher. Upon escaping the disk, these winds can enrich the so-called circumgalactic medium (CGM) (e.g. Tumlinson et al., 2011; Kacprzak et al., 2012; Werk et al., 2013), or escape the halo altogether, resulting in the enrichment of the intergalactic medium (IGM) (Davé et al., 1998; Pettini et al., 2003; Schaye et al., 2003). Recent measurements suggest that this outflowing gas is metal-enriched relative to the average ISM (Chisholm et al., 2018), further highlighting the importance of galactic winds in understanding the chemical evolution of galaxies.

Inflows of gas are important for sustaining star formation in galaxies (Erb, 2008). Enriched CGM gas may cool and rain back down on the disk, recycling ejected material (Fraternali & Binney, 2008; Marasco et al., 2012; Hobbs et al., 2015), although much of these inflows are fueled by accretion from the IGM (Sánchez Almeida et al., 2014). Inflowing gas is thus often assumed to be pristine (or at least relatively unenriched), and will dilute the metal content of galaxies. This has in particular been invoked to explain

the observed negative correlation between star formation rate (SFR) and metallicity in star-forming galaxies (e.g. Mannucci et al., 2010; Yates et al., 2012).

1.6.2 Mergers, interactions and environment

Another important consideration is the effect of mergers and interactions between galaxies. Mergers are capable of inducing starburst events in galaxies (e.g. Pearson et al., 2019). Furthermore, amalgamation of stellar mass via mergers is an important mode of galaxy growth in the hierarchical model of galaxy assembly, particularly for massive galaxies (e.g. White & Rees, 1978). Some studies suggest up to 80 % of stars in the most massive galaxies may have formed ex-situ (Oser et al., 2010), while proportions of ∼30 % have been predicted for Milky Way analogues (Pillepich et al., 2015). Mergers and strong interactions have also been observed to flatten metallicity gradients in galaxies as metal-poor gas at high galactic radius loses angular momentum and flows towards the centre (Kewley et al., 2010; Torrey et al., 2012; Jones et al., 2015a; Tissera et al., 2019).

Finally, as larger scale structure assembles, galaxies are increasingly subjected to environmental effects. Once accreted onto a larger structure (e.g. a galaxy cluster), the ability of a galaxy to replenish gas via external accretion or cooling of its CGM may be diminished ("strangulation"; e.g. Larson et al. 1980). Furthermore, depending on the density of the intracluster medium (ICM), gas from the galaxy can be stripped by ram-pressure caused by the motion of the galaxy through this medium ("gas-stripping"; Gunn & Gott 1972; Bekki 2009).

Understanding the precise balance of these processes probes physics across an incredibly vast range of physical scales: from the cosmological scale with the assembly large-scale structure, down to the atomic scale with the radiative cooling of gas. This complexity underlines not only the broad scope of the field of galaxy evolution, but also why the field generates so much interest.

1.7 Metallicity as a probe of galaxy evolution

This section will explore how measurements of the gas-phase metallicity in galaxies can be used to place constraints on the array of processes introduced in Section 1.6 that have shaped galaxy assembly. Metallicity in galaxy surveys is generally probed either in terms of global metallicities, or the spatial distribution of metals within galaxies.

1.7.1 Metallicity Scaling Relations

The most notable global metallicity scaling relation is the Mass-Metallicity Relation (MZR): the tight positive correlation between stellar mass and gas-phase metallicity observed in the local Universe over at least five orders of magnitude of mass (Lequeux et al., 1979; Skillman et al., 1989; Tremonti et al., 2004; Berg et al., 2012; Andrews & Martini, 2013; Curti et al., 2020a). The MZR is characterised by a steep increase in metallicity with stellar mass from $\sim 10^7 M_\odot$ to $\sim 10^{10} M_\odot$, with metallicity plateauing out to the highest stellar masses ($\sim 10^{12} M_\odot$).

The MZR persists qualitatively at high-redshift with an evolution toward lower normalisation (Erb et al., 2006a; Maiolino et al., 2008; Zahid et al., 2013, 2014b; Yabe et al., 2014; Sanders et al., 2020a). Some authors suggest evolution in the MZR slope to steeper values at earlier cosmic times (Maiolino et al., 2008; Zahid et al., 2013, 2014a). However Sanders et al. (2020a) find no evolution in MZR gradient out to $z \sim 3.5$ and suggest that the same physical processes contribute to the MZR slope across at least the last 12 Gyr of cosmic time. A number of the processes described in Section 1.6 can contribute to the explanation of this observed relation and its evolution with redshift.

Outflows are ubiquitous in starbursting galaxies (Heckman, 2002; Weiner et al., 2009), and there is some evidence that outflows might be metal-enriched compared to the ISM average (Chisholm et al., 2018). Since outflows are more efficient at liberating gas at lower masses, this suggests that improved metal retention at higher masses could help to explain the MZR (Tremonti et al., 2004; Chisholm et al., 2018; Sanders et al., 2020a).

Downsizing: Massive galaxies are known to evolve more rapidly at higher-redshifts than low-mass galaxies (e.g. Cowie et al., 1996). This suggests that massive galaxies will process a higher fraction of gas to yield metals at earlier times. Several authors have

reported steeper MZR measurements at earlier times suggesting that massive galaxies enrich to some empirical maximum metallicity at earlier cosmic times (Maiolino et al., 2008; Zahid et al., 2011, 2013).

Inflows: Following on from above, if low-mass galaxies are less evolved, they may preferentially sample environments where there is still significant inflow of pristine gas, which increases the gas fraction of galaxies and dilutes the metals in the ISM (Dalcanton et al., 2004; Yates et al., 2012; Bothwell et al., 2013). Several authors suggest that this plays an important role in the lower MZR normalisation observed at earlier cosmic time (Erb et al., 2006a; Sanders et al., 2020a).

Other explanations include mass-dependent **star formation efficiencies** (e.g. Brooks et al., 2007; Calura et al., 2009), or possible **variations in the stellar IMF** with galaxy mass (Köppen et al., 2007).

MZR residuals have been shown to correlate strongly with star formation rate (SFR) with galaxies having higher SFRs showing systematically lower metallicities (Ellison et al., 2008). This has been quantified as the so-called Fundamental Metallicity Relation (FMR), a 3D relation between stellar mass, metallicity and SFR (Mannucci et al., 2010; Lara-López et al., 2010). The origin of this dependence is typically explained as the balance between the inflow of pristine gas, which will promote star formation but dilute metallicity, and metal-enriched outflows (Mannucci et al., 2010; Yates et al., 2012). This has been suggested to be the manifestation of a more fundamental relationship between metallicity and the gas mass of galaxies (e.g. Bothwell et al., 2013), which correlates with SFR. In either case, some authors suggest that this relationship explains the redshift evolution of the MZR in terms of higher SFR and gas fractions at high-redshift, with measurements indicating the FMR holds out to redshifts of $z \sim 3.3$ (Mannucci et al., 2010; Richard et al., 2011; Belli et al., 2013; Stott et al., 2013; Sanders et al., 2020a), although other authors have found results that are in tension with this, instead suggesting the FMR may evolve with redshift (Yabe et al., 2012, 2014; Steidel et al., 2014; Troncoso et al., 2014; Sanders et al., 2015; Wuyts et al., 2016).

Ultimately, one of the main limitations to our understanding of the evolution of the MZR is the observational challenge associated with making reliable metallicity measurements at high-redshift spanning a wide range in mass. This is partly due to the large cosmic distances implying fainter targets. But as introduced in Section 1.5, the deeper issue

is to do with poorly constrained systematic uncertainties associated with making high-redshift metallicity measurements. This will be discussed in more detail in Chapter 3.

1.7.2 Metallicity gradients

The spatial distribution of metallicity within galaxies is a powerful tracer of the interplay of gas accretion and outflows (Edmunds & Greenhow, 1995; Magrini et al., 2016; Ma et al., 2017; Bresolin, 2019) as well as how interactions and mergers have shaped galaxies (Kewley et al., 2010; Torrey et al., 2012; Tissera et al., 2019). This is usually studied in the context of radial metallicity gradients, with metallicities in local galaxies widely observed to decrease exponentially with radius (Searle, 1971; Vila-Costas & Edmunds, 1992; Deharveng et al., 2000; Bresolin et al., 2004; Berg et al., 2013, 2020).

The advent of integral field spectroscopy has dramatically improved our ability to measure gradients in large samples of galaxies from large IFU surveys such as CALIFA (Sánchez et al., 2012), SAMI (Bryant et al., 2015) and MaNGA (Bundy et al., 2015). With the large samples available, metallicity gradient has been shown to depend on stellar mass, with low-mass galaxies having almost flat gradients and higher mass galaxies having steeper gradients (Belfiore et al., 2017; Poetrodjojo et al., 2018).

There remains some uncertainty surrounding the redshift-evolution of metallicity gradients. While some early studies found high-redshift gradients which are steeper than those observed at low-redshift (e.g. Jones et al., 2010, 2013; Yuan et al., 2011), more recent studies observe a more diverse range of gradients including many flattened gradients and even some positive gradients (Troncoso et al., 2014; Leethochawalit et al., 2016; Carton et al., 2018; Wang et al., 2019, 2020; Curti et al., 2020b; Simons et al., 2020).

If mass is considered as an evolutionary sequence (as in hierarchical assembly), flattened gradients at high-redshift are seemingly in agreement with the mass dependence observed in low-redshift surveys and paint a picture where gradients in galaxies tend to steepen over time. This is in tension with a simple picture of "inside-out" growth and instead implies that there must be strong radial mixing of gas in the early stages of stellar mass assembly, suggesting powerful feedback and outflows (e.g. Gibson et al., 2013; Hemler et al., 2020). Also, interactions and mergers have been shown to lead to a flattening of

metallicity gradients due to infall of low-metallicity gas from high-radius (Kewley et al., 2010; Torrey et al., 2012; Jones et al., 2015a).

However, within the context of the diversity of gradients observed at high-redshift, the following question is again raised: how reliable are our high-redshift metallicity measurements? In addition to being fainter, the angular resolution in observations will represent larger physical scales in high-redshift galaxies. One consequence of this is that, with access to fewer resolution elements, gradient measurements will be constructed with overall fewer data points. However the more concerning issue is the reduced ability of these low resolution observations to separate H II region emission from that of other sources. Indeed, contamination from DIG has been shown to affect gradient measurements in a complex way (Poetrodjojo et al., 2019).

The contribution of systematic uncertainties surrounding the measurement of metallicity to the scatter in metallicity gradients at high-redshift remains unclear. Improved data from future facilities will no doubt aid us in disentangling these effects. However, the continued development of our understanding of how other ionised gas properties affect metallicity measurements in galaxies will remain important. Chapter 4 will present work relevant to this topic.

1.8 Extending the parameter space of observations

In advancing our models of the chemical evolution of galaxies, idealistically we desire detailed constraints on the dark matter and baryonic components of galaxies across as broad a mass range as possible, out to the earliest epochs of galaxy evolution. Ultimately the limiting factor will always be set in part by the observational facilities at our disposal. Thus, application of observational methods at the fringes of instrumental capabilities will always have the potential to provide exciting constraints on galaxy evolution.

This thesis explores in detail a number of observational methods for measuring metallicity in galaxies and drawing connections between galaxies and their host dark matter haloes, with a particular focus on studies applied near observational detection limits. As of time of writing, the field of galaxy chemical evolution is set to be revolutionised over the coming years by new observational facilities such as *JWST* and ground-based

30 m class telescopes. It is timely, however, to remember that these facilities will not remove observational barriers, only push them deeper. Insights from observational studies presented in this thesis will contribute to maximising the impact of the exciting surveys that these new facilities will afford.

1.9 Thesis Outline

Work presented in this thesis will separately explore observational methods relating to *(i)* drawing connections between galaxies and the dark matter haloes that host them, and *(ii)* the measurement of the gas-phase metallicity of galaxies with emission lines. This work will be organised as follows.

Chapter 2 presents the application of a simple statistical method for measuring clustering of redshift $z \sim 2$ galaxies from a *HST* pure-parallel imaging survey. We expect this method to be particularly powerful when applied to future *JWST* surveys, affording powerful constraints on the connection between dark matter haloes and galaxies at redshifts beyond $z \gtrsim 6 - 8$.

Shifting focus to measuring metallicity with emission lines, Chapter 3 presents a case study exploring the prospect of extending high-redshift global strong-line metallicity measurements to lower masses with existing instruments. Chapters 4 & 5 focus on spatially resolved direct method measurements in low-redshift galaxies with auroral lines. Chapter 4 explores the effects of H II region temperature structure on direct method metallicity gradients with a galaxy from the SAMI Galaxy Survey. Chapter 5 then expands upon this, introducing some preliminary work which uses auroral emission lines detected in *Keck*/KCWI data to explore extra-planar temperature fluctuations in a low-redshift starbursting galaxy.

Finally, Chapter 6 summarises key findings presented throughout the thesis and provides a brief discussion of the implications of this work for future surveys.

Chapter 2

Observational determination of the galaxy bias from cosmic variance with a random pointing survey: clustering of $z \sim 2$ galaxies from the Hubble's BoRG survey

Gravitational clustering broadens the count-in-cells distribution of galaxies for surveys along uncorrelated (well-separated) lines of sight beyond Poisson noise. A number of methods have proposed to measure this excess "cosmic" variance to constrain the galaxy bias (i.e. the strength of clustering) independently of the two-point correlation function. Here we present an observational application of these methods using data from 141 uncorrelated fields (~ 700 arcmin^2 total) from Hubble's Brightest of Reionizing Galaxies (BoRG) survey. We use BoRG's broad-band imaging in optical and near infrared to identify $N \sim 1000$ photometric candidates at $z \sim 2$ through a combination of colour selection and photometric redshift determination, building a magnitude-limited sample with $m_{AB} \leq 24.5$ in F160W. We detect a clear excess in the variance of the galaxy number counts distribution compared to Poisson expectations, from which we estimate a galaxy bias $b \approx 3.63 \pm 0.57$. When divided by SED-fit classification into $\sim$400 early-type and $\sim$600 late-type candidates, we estimate biases of $b_{early} \approx 4.06 \pm 0.67$ and $b_{late} \approx 2.98 \pm 0.98$ respectively. These estimates are consistent with previous measurements of the bias from the two-point correlation function, and demonstrate that with $N \gtrsim 100$

sight-lines, each containing $N \gtrsim 5$ objects, the counts-in-cell analysis provides a robust measurement of the bias. This implies that the method can be applied effectively to determine clustering properties (and characteristic dark-matter halo masses) of $z \sim 6-9$ galaxies from a pure-parallel *James Webb Space Telescope* survey similar in design to Hubble's BoRG survey.

2.1 Introduction

Understanding the connection between luminous galaxies and the dark matter haloes that host them presents an important challenge in the development of galaxy evolution models. This connection encodes the lasting effect of the complex array of astrophysical processes involved in the formation and evolution of galaxies, such as star formation, tidal stripping, and mergers, that are all connected to the dynamical assembly of their parent dark-matter haloes. While direct measurements of the dark-matter halo masses are limited to cases where dynamical tracers are available at large radii (e.g. Rood, 1979; Kent & Gunn, 1982; Millington & Peach, 1986), or through gravitational lensing (e.g. Tyson et al., 1990; Dodelson, 2004), clustering measurements make it possible to draw statistical inference between the dark haloes and the properties of galaxies that inhabit them, through analytic and numerical studies of the dark-matter halo clustering (e.g. Mo & White, 1996, 2002).

Observationally, galaxy clustering is typically quantified via the construction of a two-point correlation function ($\xi(r)$), describing the excess probability of finding a galaxy at separation r from another galaxy. The galaxy bias, that is the clustering strength of galaxies relative to the matter field, relates the correlation function of galaxies ($\xi_{\rm gg}$) to that of the dark matter field ($\xi_{\rm DM}$) as a simple ratio, conventionally defined at a separation $r = 8$ Mpc.

$$b^2 \equiv \xi_{\rm gg}/\xi_{\rm DM}. \tag{2.1}$$

Thus, the observed galaxy correlation functions can be related to those derived for the dark matter field, enabling constraints to be identified on the mass of the dark matter haloes hosting these galaxies (Ouchi et al., 2004; Adelberger et al., 2005; Hayashi et al.,

2007) as well as how galaxies might be distributed within these haloes (Zehavi et al., 2005; Ouchi et al., 2005; Lee et al., 2006). Previous measurements of correlation functions conducted with large surveys such as 2dFGRS (Colless et al., 2001) and SDSS (York et al., 2000) have been successful in linking the clustering of galaxies to a range of intrinsic properties such as luminosity, colour and morphology in the local Universe and at low redshift (Norberg et al., 2001, 2002; Budavári et al., 2003; Madgwick et al., 2003; Zehavi et al., 2005, 2011; Li et al., 2006; Wang et al., 2007).

Similar measurements have been extended progressively to higher redshifts (e.g. Blanc et al., 2008; Meneux et al., 2009; McCracken et al., 2010; de la Torre et al., 2011; Wake et al., 2011; Lin et al., 2012; Mostek et al., 2013; Barone-Nugent et al., 2014; Skibba et al., 2014; Sato et al., 2014; Coil et al., 2017; Ishikawa et al., 2017; Harikane et al., 2018). Still, one key challenge for observations during the epoch of reionization ($z > 6$) is the limited area available with suitably deep infrared data, which makes legacy surveys potentially affected by systematic uncertainties due to large-scale clustering, or "cosmic variance" (Trenti & Stiavelli, 2008).

The impact of cosmic variance on high-z galaxy surveys can however be exploited to quantify the clustering properties of the galaxies, if observations span a large number of independent (uncorrelated) lines of sight, as has been innovatively proposed in a theoretical paper by Robertson (2010, R10 hereafter). The R10 method is based essentially on a counts-in-cells approach (e.g. Peebles, 1980; Efstathiou et al., 1990; Efstathiou, 1995; Andreani et al., 1994; Adelberger et al., 1998; Viironen et al., 2018), applied to a set of uncorrelated fields. In this framework, the number counts distribution will be broader than Poisson, and once the Poisson variance is removed, the remaining dispersion in the counts can be attributed to the effect of clustering, enabling a measurement of the galaxy bias.

Recently, López-Sanjuan et al. (2015, LS15 hereafter) applied a similar observational measurement to counts from the ALHAMBRA survey. In that study, the galaxy bias was successfully derived for $z \sim 0.7$ galaxies from the counts in the 48 subfields of the survey in 8 independent regions of the sky. These authors were able to compare this to bias values obtained from conventional correlation function measurements of the same survey, demonstrating consistency between the two measurements. The more sophisticated analysis employed by LS15 involves fitting a log-normal distribution to the set of number

counts. Removing the estimated Poisson shot noise from this distribution yields the intrinsic scatter from which the cosmic variance can be derived. While the work by LS15 demonstrates the feasibility of this type of analysis, the number of independent lines of sight is small.

Here, we address this limitation and use a large *Hubble Space Telescope* (*HST*) random pointing survey, the Brightest of Reionizing Galaxies (BoRG) survey (Trenti et al., 2011; Bradley et al., 2012; Schmidt et al., 2014; Calvi et al., 2016) to assess the performance of the LS15 and R10 bias measurement methods using over 140 independent lines of sight. BoRG has been designed to identify rare bright ($L > L_*$) galaxies at $z \gtrsim 7$ from broad-band imaging, minimizing the effects of cosmic variance on the galaxy luminosity function. While the number counts of these rare sources do not allow an application of the method (Poisson noise dominates when the average number of counts per field is below unity), the $VYJH$ filter set of BoRG also proves ideal for the identification of $z \sim 2$ galaxies via the 4000 Å break. Thus, in this paper we apply these methods to the counts of photometrically selected $z \sim 2$ galaxies from BoRG to estimate their bias, and investigate how it depends on galaxy properties such as spectral type to demonstrate the efficiency of the technique and compare its results against those from the literature based on the two point correlation function.

The paper is structured as follows. In section 2.2 we introduce the BoRG survey and give an overview of the data we have used including some of the technical aspects. Section 2.3 outlines the methods we have used in creating source catalogs and selecting $z \sim 2$ candidates, including the subtle differences between the two BoRG sub-samples, as well as how the LS15 and R10 methods are applied to our number counts to calculate the large-scale galaxy bias. Section 2.4 outlines the results of our analysis. In sections 2.4.1 and 2.4.2 we obtain a preliminary bias values for the BoRG[z8] and BoRG[z9] sub-samples considered separately. Section 2.4.3 then details our final bias values and how they compare to previous measurements. We summarise in Section 2.5, and discuss applications to future observations of high-z galaxies with the *James Webb Space Telescope*.

In this paper we adopt the Planck Collaboration et al. (2016) cosmology: $\Omega_\Lambda = 0.692$, $\Omega_M = 0.308$, $\sigma_8 = 0.815$ and $H_0 = 67.8$ km s^{-1} Mpc^{-1}. All magnitudes are quoted in the AB magnitude system (Oke & Gunn, 1983).

2.2 Data

The Brightest of Reionizing Galaxies (BoRG) survey (Trenti et al., 2011) is an optical to near-infrared multi-waveband pure-parallel survey conducted with Wide-Field Camera 3 (WFC3) on *HST* with the goal of finding the most luminous galaxies at redshifts $z \gtrsim 7$ using the Lyman-break dropout technique (Steidel et al., 1996).

The survey consists of two generations spanning a number of *HST* cycles, with the focus shifting from identification of Lyman-break dropout candidates at $z \sim 8$ in the first generation, BoRG[z8] (Trenti et al., 2011), to pushing this search to *HST*'s detection limit at $z \sim 9 - 10$ in the second generation, BoRG[z9] (Calvi et al., 2016).

The BoRG[z8] data set consists of 71 independent pointings imaged with WFC3/IR in three different near-infrared filters (F098M, F125W and F160W) with additional imaging conducted with WFC3/UVIS in one optical filter (either F606W or F600LP). In this paper we use the 2014 (DR3) public release[1].

Additionally, we use data collected with BoRG[z9], consisting of the initial 28 fields presented in Calvi et al. (2016) as well as 46 new pointings imaged with WFC3/IR in four different near-infrared filters (F105W, F125W, F140W and F160W) with optical imaging conducted with WFC3/UVIS in the F350LP filter.

The design of pure-parallel surveys is limited by constraints imposed by the primary observations, with exposure times varying in each available opportunity. This leads to a non-uniform survey, with 5σ limiting magnitudes (for point sources and aperture $r = 0''.2$) ranging between $m_{\rm AB} = 25.6 - 27.5$. Also, the data do not include standard dithering (not employed by spectroscopic primary programs). However, a combination of careful planning to take advantage of natural roll-angle dithering between different orbits and of Laplacian edge detection filtering allows us to effectively identify and remove artifacts, such as hot/cold pixels that have not been flagged by the calibration pipeline. Overall, the image quality is very close to prime dithered observations (see Calvi et al. 2016 for a comprehensive discussion). For more detailed discussions of the technical details of the BoRG survey, the reader is referred to Trenti et al. (2011, 2012); Bradley et al. (2012); Schmidt et al. (2014); Calvi et al. (2016).

[1] https://archive.stsci.edu/prepds/borg/

Two fields, borg_0751+2917 and borg_1209+4543, were serendipitously imaged in both the BoRG[z8] and BoRG[z9] data sets, due to the same primary target being revisited during multiple cycles. To avoid counting the same fields twice, only the BoRG[z8] pointings were considered for these fields in our galaxy bias analysis. However, these fields presented a unique opportunity to compare the photometry of the two separate filter sets considered in our analysis. This comparison is discussed in section 2.3.2.3.

From the remaining sample, one field was discarded due to a problematic over-density of stars which was covering the majority of the area in the image, while one was discarded due to significant Galactic dust-reddening which could have affected the photometric redshift selection if small scale fluctuations in reddening had been present. This left us with a combined sample of 141 uncorrelated fields spanning $\sim$700 arcmin^2.

2.3 Methods

In this analysis, we identify $z \sim 2$ candidates with a $Y - H$ colour selection, capturing the prominent break at 4000 Å in rest-frame luminosity, and subsequent photometric redshift estimation. The importance this places on the Y band photometry leads us to initially consider the BoRG[z8] and BoRG[z9] data sets separately as different Y filters are employed (F098M vs. F105W). This section introduces the analysis methods generally and describes specific application to the BoRG[z8] sample before outlining how the difference in Y band filter is mitigated.

2.3.1 Source Catalogs

The BoRG science data consist of reduced and aligned science images produced with `MultiDrizzle` (Koekemoer et al., 2003) as well as associated weight maps. From each weight map we create corresponding rms maps. In addition, a super-rms map is created by flagging all pixels that have zero weight in at least one of the filters. In this way we can include in the photometric analysis only pixels with complete spectral coverage, minimizing the impact of detector artifacts in the source catalogs and their associated photometric measurements.

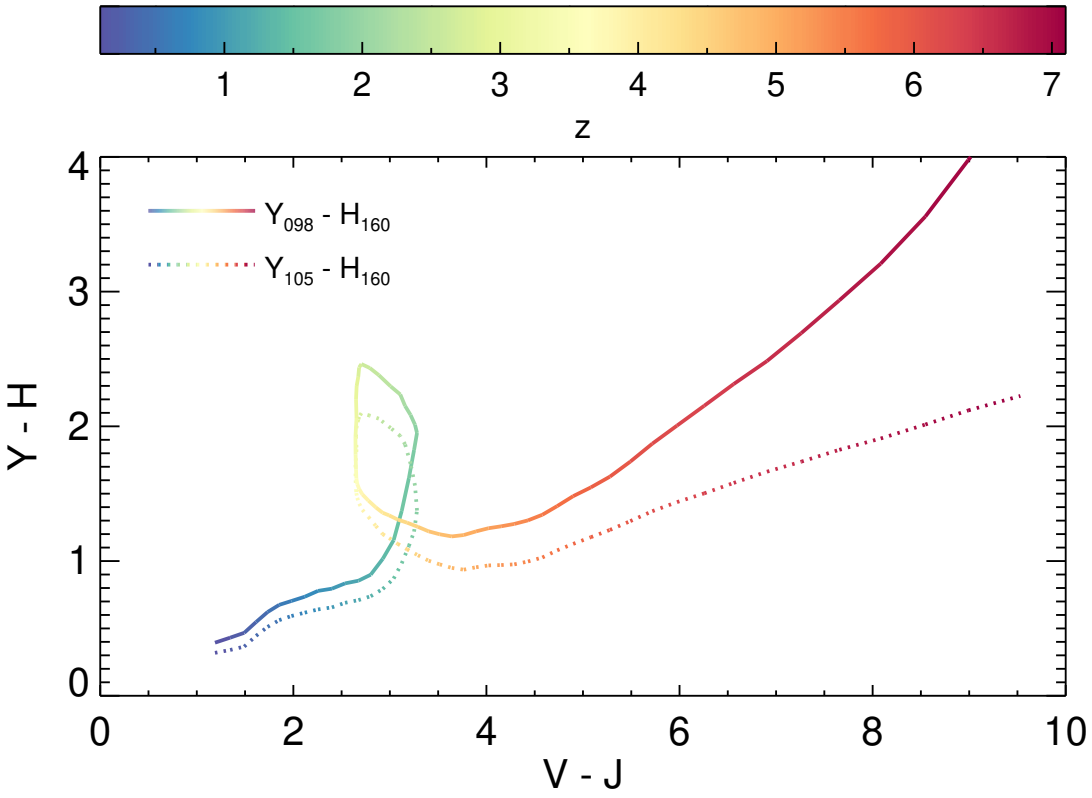

FIGURE 2.1: Plot of the $Y - H$ and $V - J$ colours changing with redshift for an old passive galaxy for filter sets with Y_{098} and Y_{105} filters. $Y_{098} - H_{160} > 1.5$ colour cut is an important preselection in the identification of $z \sim 2$ candidates from the $YHVJ$ filter set used in the BoRG survey. $V - J$ colours are not directly used in our $z \sim 2$ candidate selection.

The rms maps are then normalised to account for correlated noise induced by `MultiDrizzle` (Casertano et al., 2000). This is achieved by measuring the noise in randomly placed empty apertures (i.e., no overlap with detected sources) and comparing this to the noise value inferred from the rms map. Any disparity is corrected by applying a scaling factor to the rms map (see Trenti et al. 2011).

Photometric zeropoints are set individually for each field by correcting the default *HST* zeropoints for Galactic reddening along the line of sight, estimated from the $E(B - V)$ reddening factors in Schlafly & Finkbeiner (2011).

Source catalogs in each filter were created for each field using `SExtractor` (Bertin & Arnouts, 1996a) in dual-image mode with source detection performed in the H_{160} band, with further pruning to include only sources with $S/N \geq 5$ in the detection image, where signal-to-noise ratios were calculated using isophotal flux and error (`FLUX_ISO`/`FLUXERR_ISO`). The `SExtractor` `CLASS_STAR` parameter was used to exclude stars from the analysis, with sources having `CLASS_STAR` > 0.8 being excluded. Given the non-uniform depth of the

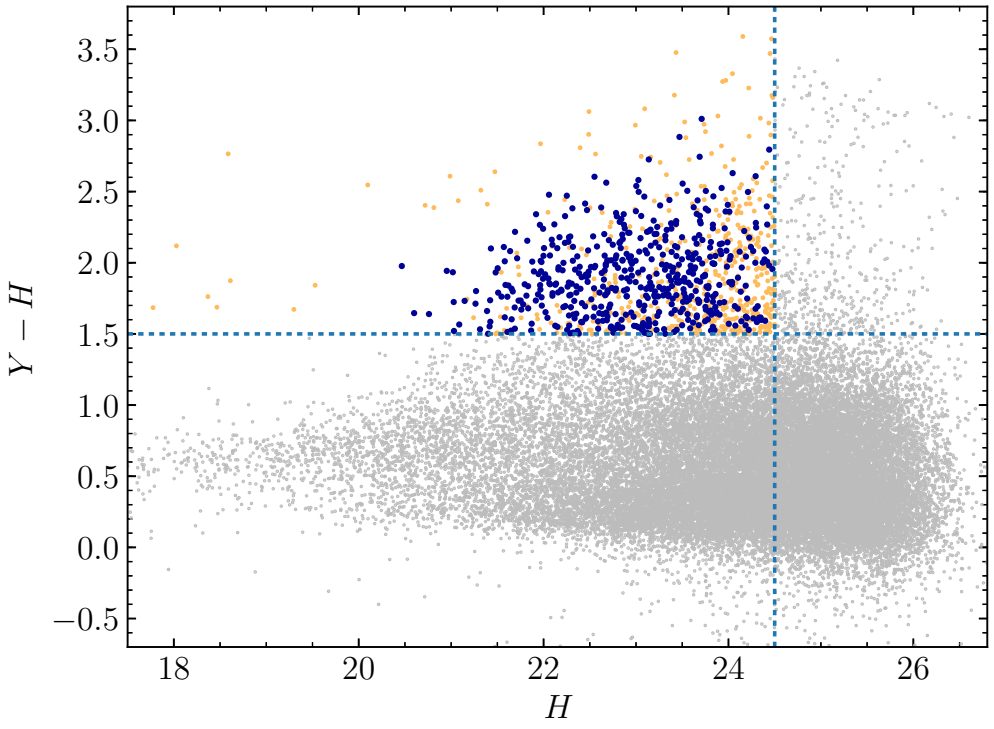

FIGURE 2.2: Colour–magnitude plot of all resolved sources in the BoRG[z8] sample. Dotted lines show magnitude cut-off ($H < 24.5$) and colour cut ($Y - H > 1.5$). Blue and yellow markers show sources passing these selection criterion. Blue markers depict visually confirmed candidates with photometric redshifts that passed all criteria and were thus included in our final sample ($1.5 < z < 2.5$, `chisq1` < 1 and `odds` > 0.8), while the yellow markers depict sources that failed one or more of these criteria and were excluded from our sample. Grey markers show all remaining sources detected with $S/N > 8$ that did not meet either or both of the magnitude cut-off or the $Y - H$ colour cut.

BoRG survey, we restricted the analysis to sources with `MAG_AUTO` < 24.5, which ensures a magnitude-limited sample with high completeness (despite the fact that several fields are complete to fainter magnitudes).

2.3.2 Redshift $z \sim 2$ Candidate Selection

Near-infrared imaging with WFC3/IR using a $YHVz$ filter set has been shown to enable effective photometric selection of galaxies in the redshift range $1.5 < z < 3.5$. Cameron et al. (2011) devised a $Y - H$ versus $V - z$ colour-colour selection to select for passive and star-forming galaxies in HUDF and GOODS-S, exploiting the presence of the rest-frame 4000 Å break in this wavelength range. Their selection involves two criteria: star-forming galaxies are captured by selecting sources redder in $Y - H$ colour than in

$V - z$ by a given amount, while the passive population (very red in $V - z$) is captured with a simple cut, selecting all sources with $Y - H$ colour redder than a given value, determined by the Y-band filter used.

The limited filter set available in the BoRG survey compared to a legacy field survey prompted us to employ a one-colour selection followed by full application of photometric redshift to build our final sample. Figure 2.1 shows the progression of $Y - H$ and $V - J$ colours of an example passive galaxy spectrum as it is moved to higher redshift. The knot in figure 2.1 around redshift $z \sim 2$, caused by the 4000 Å break lying between the Y and H bands, is what we aim to capture with this $Y - H$ one-colour selection.

2.3.2.1 BoRG[z8] candidate selection

For the Y_{098} filter used in the BoRG[z8] sample, we apply a cut of $Y_{098} - H_{160} > 1.5$, motivated by the work of Cameron et al. (2011), to select for passive galaxies as well as dust-reddened star-forming galaxies. Photometric redshift estimates were then obtained for sources passing this colour selection with `BPZ` (Benítez, 2000; Benítez et al., 2004; Coe et al., 2006), a Bayesian Photometric Redshift software. Input catalogs were constructed for `BPZ` by compiling the isophotal magnitude (and the associated rms error) in each filter for each source (`MAG_ISO` and `MAGERR_ISO`).

Our final sample of $z \sim 2$ candidates was selected based on parameters from the resultant `BPZ` output catalogs. Sources in the redshift range $1.5 < z < 2.5$ were selected based on the photometric redshift best-estimate (`zb`) output parameter. To ensure quality-of-fit for the redshift estimates, sources with `chisq2` > 1 were eliminated. Additionally, the `odds` measure (likelihood that `zb` is correct within 0.10) was used to exclude sources with multiple possible redshift solutions, with sources having `odds` < 0.8 being removed.

As a final vetting, all candidates were visually inspected and removed if they were judged to be photometric artifacts. From the initial 987 candidates across the full survey (BoRG[z8] and BoRG[z9]), 18 sources were removed due to non-equivocal identification as part of diffraction spikes created by bright stars. The colour-magnitude diagram in figure 2.2 shows the location of resolved sources passing the set of selection criteria in the BoRG[z8] sample.

2.3.2.2 BoRG[z9] candidate selection: $Y_{105} - H_{160}$ Colour Cut

The BoRG[z9] sample consists of a further 74 fields imaged on WFC3 with a slightly different filter set (see section 2.2). In particular, the BoRG[z9] sample employs a different Y-band filter, the wide band F105W filter, rather than the medium band F098M filter used in BoRG[z8]. Here we describe our approach to selecting a new $Y - H$ colour cut to mirror this change in filters. Other than this alternative colour cut, the same method was applied to the BoRG[z9] sample as for BoRG[z8], described in section 2.3.2.1. An overview of the results for the BoRG[z9] sample, considered in isolation to BoRG[z8], is presented in section 2.4.2.

The efficiency of source selection and resultant sample redshift range depends strongly on the $Y - H$ cut used. Thus, the alternative Y_{105} filter used for the BoRG[z9] sample requires a different magnitude colour cut to be applied. It is expected that a slightly lower cut-off would be required as the redder average wavelength of the F105W filter compared to the F098M filter would imply a particular $z \sim 2$ galaxy would appear bluer in $Y_{105} - H_{160}$ colour than in $Y_{098} - H_{160}$.

We considered two metrics for comparing the BoRG[z9] colour cut to that used in BoRG[z8]. Increasing the cut-off colour value would decrease mean number counts, but should increase the mean redshift derived for the population, due to the increased emphasis on redder objects. We reasoned that both of these quantities should be consistent between the two samples as they reflect properties of the galaxy populations being sampled. Figure 2.3 summarises our investigation of how different $Y_{105} - H_{160}$ cuts impacted the mean counts per field and redshift distribution of the resultant sample.

Given the importance of Poisson statistics to the analysis, we eventually selected the cut that best reproduced the mean counts seen in BoRG[z8], adopting a cut-off of $Y_{105} - H_{160} > 1.2$. It is perhaps worth noting that although in varying this cut we influence the first moment of the distribution of counts (the mean), the derived galaxy bias depends primarily on the second moment (variance).

2.3.2.3 Colour cut comparisons with borg_0751+2917 and borg_1209+4543

The fields borg_0751+2917 and borg_1209+4543 were serendipitously imaged in both BoRG[z8] and BoRG[z9], observed in parallel to separate revisits to the same primary

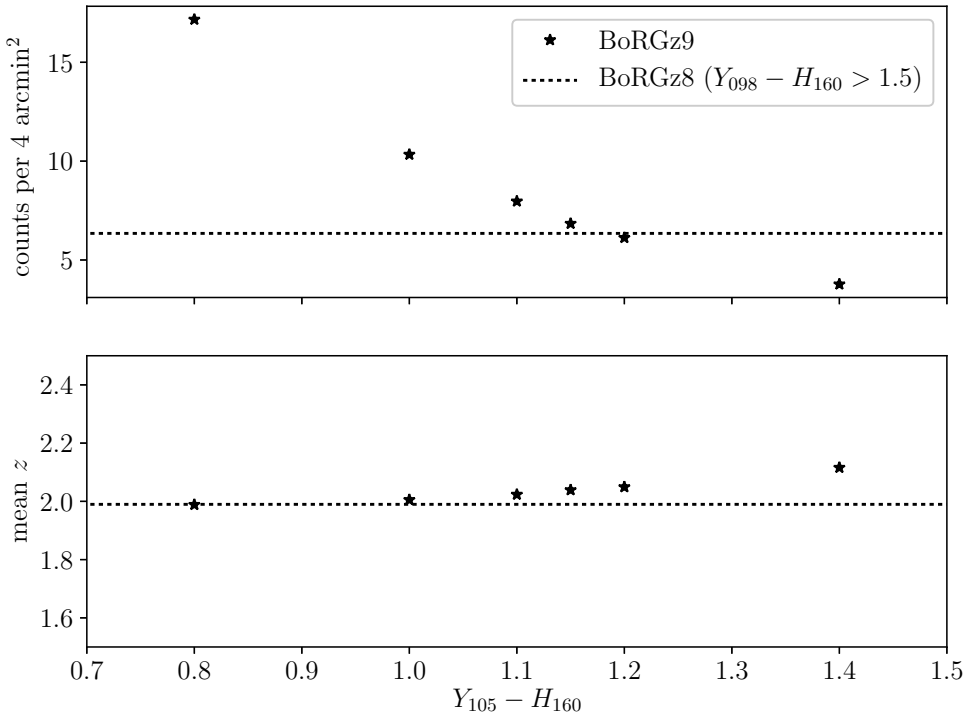

FIGURE 2.3: Mean scaled field counts and mean photometric redshift recovered from BoRG[z9] analysis with different $Y_{105} - H_{160}$ colour cuts. Horizontal dotted lines show benchmark values obtained from BoRG[z8] analysis using $Y_{098} - H_{160} > 1.5$. It is worth noting that these numbers include sources that would eventually be removed as artifacts by visual inspection. Given that only $\sim 1\%$ of sources would be excluded on these grounds, it was considered that this would not impact the outcome of this comparison.

target. Though in both cases the BoRG[z9] visits were discarded for our final analysis so as not to consider the same galaxy candidates twice in our set of field counts, they present a unique opportunity to compare the $Y - H$ colour cuts selected for each sample.

The $z \sim 2$ candidate counts returned for the borg_0751+2917 were eight and nine for the BoRG[z8] and BoRG[z9] samples respectively. The eight candidates counted in BoRG[z8] were all present in the BoRG[z9] final count along with one additional candidate, borgz9_0751+2917_317. In that case, the candidate in the BoRG[z8] pointing had been positioned near the edge of the field, affecting the detected shape of the object, ultimately resulting in the object not passing colour selection, with a colour of $Y_{098} - H_{160} = 1.44$.

The BoRG[z8] and BoRG[z9] samples both returned six $z \sim 2$ candidates from the borg_1209+4543 field. Four of these candidates were observed in both fields with each

sample returning two candidates not observed in the other. Of these sources selected in one sample but not the other, one was near the $m_{AB} = 24.5$ threshold and was excluded in BoRG[z8] based on its magnitude, but selected in BoRG[z9]. The other three candidates were excluded from one sample based on the $Y - H$ colour cut, but selected in the other. In all three of these cases, the colour cut-off was well within the 1σ uncertainty of the $Y - H$ value (given by `SExtractor`) for one of either the BoRG[z8] or the BoRG[z9] sample. In this sense, the fact that these sources were selected as $z \sim 2$ candidates in one sample and not the other can be attributed to uncertainty in the photometry that is within the bounds of what is expected.

Between the two fields, a total of 14 sources passed colour selection in both samples. Among these sources twelve were selected as $z \sim 2$ candidates by both samples and two were rejected by both samples (*i.e.*, the two samples were in agreement in all cases). Of the twelve that were selected as $z \sim 2$ candidates, the photometric redshifts obtained from each sample were in fairly good agreement. The largest photometric redshift discrepancy between the two samples was $|z_{\mathrm{BoRG[z8]}} - z_{\mathrm{BoRG[z9]}}| = 0.33$, with only four such discrepancies in excess of 0.20. The two rejected sources were excluded by both samples based on a low `odds` value, indicating the photometry was unable to distinguish between multiple possible redshift solutions in both cases.

These two fields unfortunately only make up a very small sample with which to compare how the selected colour cuts and photometry perform for the BoRG[z8] and BoRG[z9] samples. However, the limited crossover we have suggests that, within the bounds of the uncertainties of photometry for objects at this redshift, the two samples with the colour cuts as selected are yielding comparable galaxy populations.

2.3.3 Normalization of Field Effective Area

An accurate and unbiased comparison of number densities of $z \sim 2$ galaxies in different fields is critical to this analysis, however differences between the LS15 and R10 methods mean that the normalization is implemented in different ways (these methods are outlined in more detail in section 2.3.4). Key to this normalization is an evaluation of the effective area sampled by each field. Several fields from the survey are combinations of two or more pointings and hence the angular sizes of these fields take a range of values. Moreover, large foreground objects may mask many $z \sim 2$ sources from detection.

The presence of highly extended low-redshift galaxies and diffraction spikes from bright galactic stars must therefore also be taken into account.

For each field, an effective area was calculated from the `MultiDrizzle` weight maps and `SExtractor` segmentation check images to enable this comparison. From the weight map we summed the number of pixels that received a non-zero exposure time. These pixels were then excluded if they returned a non-zero value in the segmentation image, thus removing any pixels that were part of a detected object. Consequently, the presence of large objects obscuring a significant proportion of a field would be reflected in the derived area.

In the LS15 method, the different fields are considered as volumes of varying size and the number counts used are the raw integer number counts obtained as per section 2.3.2.1. The co-moving volumes were calculated with the online calculator available from Trenti & Stiavelli (2008) as described in section 2.3.5.

However, the R10 method requires us to compare densities as though the fields sampled identical volumes. Thus for this method, the number counts for each field were scaled according to these areas such that they reflect a representative sampling of the galaxy density at that location in the sky. Additionally, key to the R10 method for estimating galaxy bias is removal of the scatter due to Poisson variance from the distribution of galaxy counts. Due to the dependence of Poisson variance on the mean, it is important that scaling these galaxy counts preserves a mean that is representative of the survey sample volume. Thus, a scaling factor was applied to each field such that its count reflects an area of 4.41 arcmin^2, the median area calculated for the survey. This set of values was used in the calculation of the galaxy bias. The full field-by-field number counts can be found in Appendix A.2.

2.3.4 Bias Calculation

Galaxy bias is the ratio of the two-point correlation functions of galaxies and matter (equation 2.1). Assuming the bias in equation 2.1 is scale-independent, the variance of galaxy (σ_g^2, or "cosmic variance") and dark matter (σ_{DM}^2) fields is related to the bias through

$$\sigma_g^2 = b^2 \sigma_{DM}^2. \tag{2.2}$$

We employ two independent methods to obtain σ_g^2. The key quantity employed by both methods is the intrinsic scatter in a set of number counts from a large number ($N \gtrsim 50$) uncorrelated fields that represent discrete samplings from the underlying galaxy density field.

In sections 2.3.1 – 2.3.3 we obtained a set of field-by-field $z \sim 2$ candidate counts. The variance of this set, σ_N^2, is expected to have two components: one due to the intrinsic clustering on large scales ($\sim \sigma_{\rm g}^2$), and one due to the Poisson statistics associated with taking discrete samplings from the galaxy density field.

The two methods differ subtly in their approaches toward disentangling these effects. The different methods are described in the following sections (2.3.4.2 - 2.3.4.1)

2.3.4.1 Robertson Method

The Robertson (2010) method is based in the assumption that in the absence of clustering, our galaxy counts would fit a Poisson distribution with $\sigma_{\rm Poisson}^2 = \bar{N}$. In fact, our counts exhibit an excess scatter, with $\sigma_N^2 > \bar{N}$, attributed to large scale clustering (see Figure 2.5). Following R10, we calculate this excess to yield $\sigma_{\rm g}^2$,

$$\sigma_{\rm g}^2 \approx \frac{\sigma_N^2 - \bar{N}}{\bar{N}^2}. \tag{2.3}$$

Given we expect σ_g^2 and σ_{DM}^2 to be related linearly via the bias, we can use this value of σ_g^2 and a value of σ_{DM}^2 predicted from theory (see section 2.3.5) to estimate the galaxy bias.

$$b^2 = \frac{\sigma_g^2}{\sigma_{DM}^2} \approx \frac{\sigma_N^2 - \bar{N}}{\bar{N}^2 \sigma_{DM}^2}. \tag{2.4}$$

2.3.4.2 López-Sanjuan Method

The López-Sanjuan et al. (2015) method differs in that it obtains a measure of the intrinsic scatter by fitting a log-normal distribution to the number densities after estimating the Poisson shot noise for each volume.

The variance in galaxy counts due to cosmic variance is defined as

$$\sigma_{\mathrm{g}}^2 = \frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle^2} - \frac{1}{\langle N \rangle}. \tag{2.5}$$

The $1/\langle N \rangle$ term is to correct for Poisson noise. In the LS15 method, Poisson shot noise is estimated to be

$$\sigma_{P,j} = \sqrt{N_j}/V_j \tag{2.6}$$

where N_j and V_j are the number counts and volumes of sub-field j. Using the maximum likelihood estimator (MLE) presented in López-Sanjuan et al. (2014), a log-normal distribution is fit to our field densities (number counts obtained in section 2.3.2.1 scaled according to co-moving volumes derived in line with section 2.3.5) and the median number density $\bar{n}$ and intrinsic dispersion σ_{int} for the sample are obtained.

Applying equation 2.5 to $P_{\mathrm{LN}}(n)$, LS15 conclude that the observed relative cosmic variance is

$$\sigma_g^2 = e^{\sigma_{\mathrm{int}}^2} - 1. \tag{2.7}$$

We then obtained a value for galaxy bias b_g by applying equation 2.2 using a value for σ_{DM}^2 obtained as described in section 2.3.5.

2.3.5 Value for variance of the matter field (σ_{DM}^2)

In calculating the galaxy bias from the methods described, we consider σ_{DM}^2 over the volume spanned by the median field area (2.1 arcmin $\times$ 2.1 arcmin) across the redshift interval $\bar{z} - \sigma_z < z < \bar{z} + \sigma_z$, where $\bar{z}$ and σ_z are the mean and standard deviation of the

TABLE 2.1: Values used for variance of matter field in bias calculation. All values are calculated with Trenti & Stiavelli (2008) using the available online calculator. Comoving dimensions of the pencil beam geometry adopted are given. A value of $\sigma_8 = 0.815$ is assumed (Planck Collaboration et al., 2016).

Sample	z interval	σ^2_{DM}	Comoving Dimensions Mpc h^{-1}
Combined	2.02 ± 0.26	0.0118	$2.31 \times 2.31 \times 566$
BoRG[z8]	1.99 ± 0.25	0.0125	$2.29 \times 2.29 \times 549$
BoRG[z9]	2.05 ± 0.27	0.0112	$2.23 \times 2.23 \times 574$
Early type	1.89 ± 0.27	0.0120	$2.22 \times 2.22 \times 615$
Late type	2.11 ± 0.22	0.0139	$2.37 \times 2.37 \times 452$

photometric redshifts obtained across the whole sample. This corresponds to comoving dimensions of approximately $2.3 \times 2.3 \times 560$ Mpc$^3 h^{-3}$ for the BoRG[z8] sample (assuming a field of view of 4.41 arcmin^2).

We calculate σ^2_{DM} using the online calculator available from Trenti & Stiavelli (2008), assuming $\sigma_8 = 0.815$ (Planck Collaboration et al., 2016). We recalculate a new value for each survey subsample considered, using the precise redshift distribution obtained for that subsample. A summary of the volumes considered and values obtained for each of the subsamples in our analysis can be found in Table 2.1.

2.3.6 Bootstrapping Error Analysis

The López-Sanjuan et al. (2014) MLE used for the LS15 method provides 1σ error estimates. For values calculated from the R10 method, we quantify the uncertainty with a Monte-Carlo bootstrapping resampling method. In this method, counts from the original data set are selected at random with repetition allowed to construct a synthetic data set of length N_{fields} with which a new bias calculation could be performed (section 2.3.4). This bias calculation is repeated 1000 times for newly generated synthetic sets, yielding a distribution of bias values. The 1σ uncertainty of this distribution was adopted as the 1σ uncertainty for the calculated bias value.

2.4 Results and analysis

In this section we describe the candidate sample obtained as well as the bias values obtained from each of the LS15 and R10 methods. Since candidates in the BoRG[z9]

sample were selected using a different colour cut (see section 2.3.2.2), we initially consider each of the BoRG[z8] and BoRG[z9] samples in isolation (sections 2.4.1 and 2.4.2). After establishing the bias results are broadly comparable in both samples (as expected), we combine the two samples and compute overall bias values with both methods (section 2.4.3.2). We then divide the candidates into early-type and late-type sub-samples to examine the dependence of the bias on spectral type (2.4.3.3). It is worth noting that the two methods give different values, with disagreements in excess of 1σ in some cases. We discuss the differences between the two methods in section 2.4.4.

2.4.1 BoRG[z8] Sample

We identify 490 $z \sim 2$ candidates across the 69 fields in the BoRG[z8] survey with a mean count of 7.1 candidates per field and a field-to-field variance of 13.1.

All but the most luminous candidate have H-band magnitudes of $20.5 < m_H < 24.5$, with the brightest candidate having $m_H = 20.47$. Number counts increase steadily with decreasing luminosity to a magnitude of $m_H = 23.0$ with $23.0 < m_H < 23.5$ being the most populated magnitude bin. A higher proportion of sources failing photometric redshift selection at magnitudes fainter than $m_H = 23.5$ leads to a decrease in number counts out to $m_H = 24.5$, despite a proportionally high number of sources passing $Y - H$ colour selection across this magnitude range.

Approximately 60% of candidates passing all selection criteria had S/N in the range $5 < S/N < 50$, while only 9% had $S/N > 100$. Sources for which we obtained photometric redshifts were selected to fit the range $1.5 <$ `zb` < 2.5 with a resultant mean redshift of $\bar{z} = 1.99$ and standard deviation of $\sigma_z = 0.25$ across the sample, as seen in figure 2.4. Quality-of-fit selection was conducted based on `odds` and `chisq2` output parameters from `BPZ`. The criteria `odds` > 0.8 and `chisq2` < 1.0 were selected for with 86% of candidates having `odds` > 0.9 and 82% with `chisq2` < 0.50.

Applying the methods described in sections 2.3.4 – 2.3.6 to the set of counts obtained from the BoRG[z8] sample yields a preliminary value for the galaxy bias of $b \approx 2.88 \pm 0.60$ derived from the LS15 method or $b \approx 3.20 \pm 0.68$ as derived from the R10 method (table 2.2; refer to section 2.4.4 for a discussion of the difference in values obtained from the

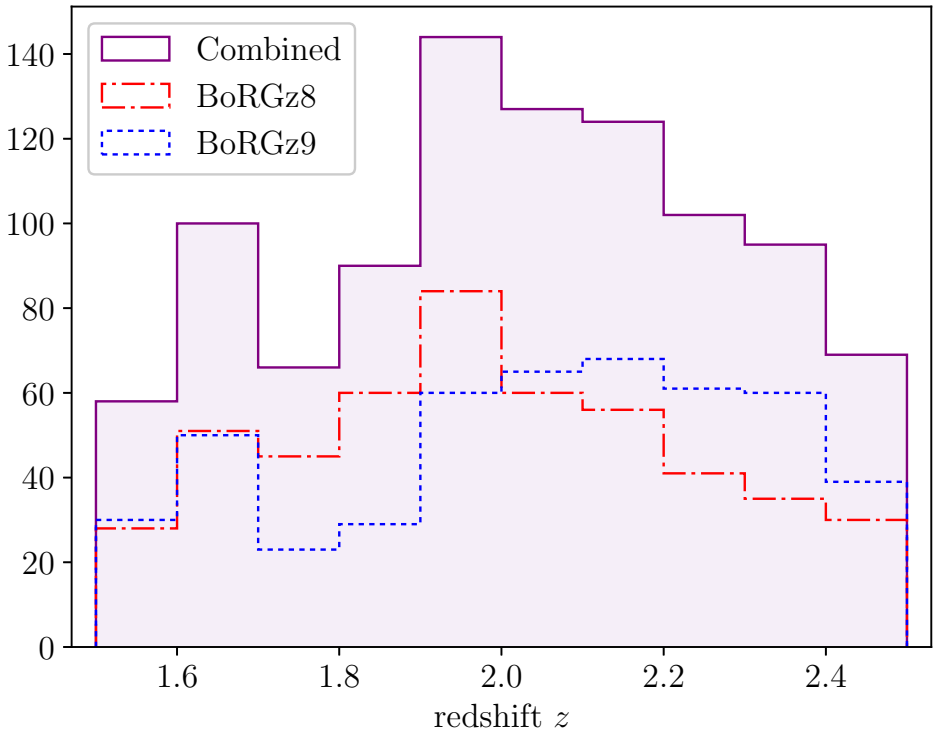

FIGURE 2.4: Distribution of best photometric redshifts as determined with `BPZ` of $z \sim 2$ candidates passing colour cut and `BPZ` quality-of-fit criteria. Dash-dot line shows the initial BoRG[z8] sample with the $Y_{098} - H_{160} > 1.5$ colour cut, while dotted line shows the BoRG[z9] sample using a $Y_{105} - H_{160} > 1.2$ colour cut. The solid line depicts the final redshift distribution after combining the two sets.

two methods). Figure 2.5 shows histograms of the number of fields with given numbers of candidates.

2.4.2 BoRG[z9] Sample

The BoRG[z9] sample consists of 74 fields, including the 2 fields identified as BoRG[z8] revisits (see section 2.3.2.3), totalling 72 new unique fields. We identify 494 $z \sim 2$ candidates across these 74 fields at a slightly lower mean of 6.7 candidates per field, but with a higher field-to-field variance of 14.9.

The BoRG[z9] sample shows no significant deviations from BoRG[z8] in terms of the magnitude and signal-to-noise ratio distributions it returns, with the brightest candidate having a magnitude of $m_H = 20.81$.

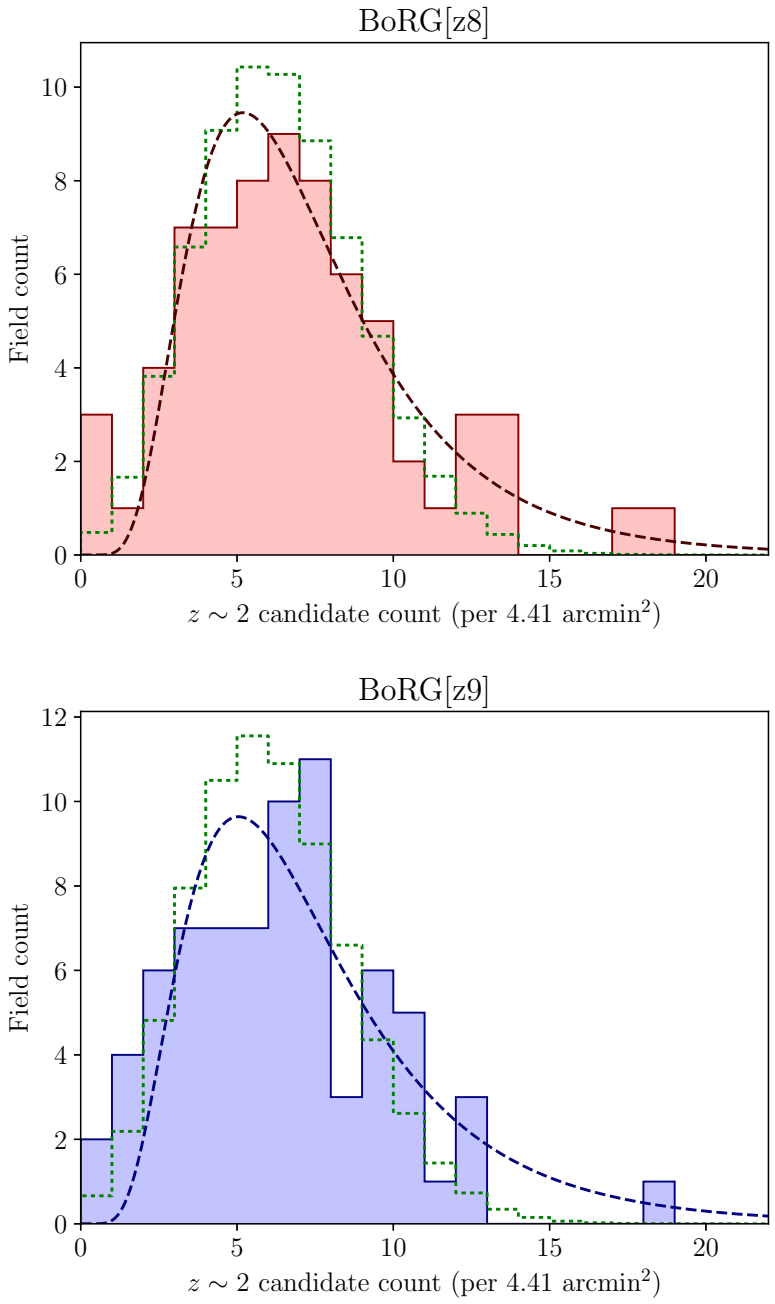

FIGURE 2.5: Distributions of $z \sim 2$ candidate counts across fields in the BoRG[z8] and BoRG[z9] subsamples of the BoRG survey. The scatter in these distributions is expected to have two components: one due to Poisson uncertainty and one due to cosmic variance. The green dotted histogram shows a Poisson distribution with the same mean and normalisation as the set of field counts. In both the BoRG[z8] and BoRG[z9] samples, a broadening of set of field counts beyond Poisson uncertainty is observed, as expected. The R10 method derives an estimate of the galaxy bias from this simple excess variance. The LS15 method fits a log-normal distribution to the densities (dashed line shows log-normal fit to scaled field counts for illustration). Note that the plotted line is not the distribution used for calculating the bias, that involves fitting a normal to log densities. The intrinsic scatter, or "cosmic variance" is derived from the log-normal fit, affording a value for the galaxy bias. These methods are described in more detail in section 2.3.4.

Photometric redshifts from `BPZ` returned a mean redshift of $\bar{z} = 2.05$ and a standard deviation of $\sigma_z = 0.27$, both values slightly higher than those from BoRG[z8]. The distribution of quality-of-fit (`odds` and `chisq2`) values showed no significant change from that seen in BoRG[z8].

We calculated a value for the galaxy bias for BoRG[z9] considered in isolation as a final comparison of the candidate populations of the two samples. Following the procedure outlined in section 2.3.4 – 2.3.6, we obtained a new σ^2_{DM} value, to reflect the subtly different redshift distribution observed, before obtaining a galaxy bias of $b \approx 2.93 \pm 0.61$ from the LS15 method and $b \approx 3.94 \pm 1.02$ from the R10 method. A slightly higher field-to-field variance in the scaled counts in the BoRG[z9] sample results in a higher estimate for the bias when compared to the BoRG[z8] results with both methods. However, both methods yield values that are within 1σ uncertainty of the BoRG[z8] value. Encouragingly, the uncertainties for the two values derived from the LS15 method are very similar, as expected due to a similar number of fields being present in each sample. These similar values further suggest to us that the two samples are returning comparable candidate populations.

2.4.3 Combined Sample

2.4.3.1 Overview

In order to implement the two methods on as large a data set as possible, we desired a combined sample made up of number counts from fields across both the BoRG[z8] and BoRG[z9] samples. The selection process for $z \sim 2$ candidates is identical in both samples, save for different cutoff values in the $Y - H$ colour cut as discussed in sections 2.3.2.2 – 2.3.2.3. Satisfied that these two colour cut values were returning comparable candidate populations, we combined the sets of field counts obtained for each sample into one combined set with which we conduct our final bias analysis.

This yielded a final data set covering 141 fields (we ignore the BoRG[z9] visits to the two doubly imaged fields discussed in section 2.3.2.3). From these fields our sample consists of a total of 969 candidates, with a mean of 6.9 candidates per field and a variance of 14.3 across the complete sample.

The distribution of photometric redshifts obtained across the complete sample had a mean of $\bar{z} = 2.02$ and a standard deviation of $\sigma_z = 0.26$. The distribution of magnitude, signal-to-noise ratio, and photometric redshift quality-of-fit diagnostics among the candidates show no significant deviation from those described for the BoRG[z8] sample in section 2.4.1. The distribution of field counts obtained for the combined sample is shown in figure 2.6.

2.4.3.2 Overall bias value

Applying the LS15 method (see sections 2.3.4 – 2.3.6) to the combined set of field counts yielded a value of $b \approx 2.94 \pm 0.41$ for our complete sample, while the R10 method arrived at a value of $b \approx 3.63 \pm 0.57$ (table 2.2; probability distribution of the R10 bias value obtained with Monte-Carlo bootstrapping is shown in figure 2.7). No previous studies have measured the bias of galaxies at $z \sim 2$ using the same selection criteria as applied in this analysis, making direct comparisons to the literature somewhat difficult. Wake et al. (2011) explored the evolution of bias of galaxies of different stellar masses between $z \sim 1 - 2$ using photometric redshift estimations applied to the NEWFIRM Medium Band Survey. The bias for the lowest stellar mass bin of the $\bar{z} = 1.9$ sample identified in this survey was measured at $3.30^{+0.45}_{-0.23}$, broadly in line with our measurements. However, this corresponds to the NMBS magnitude limit of $K = 22.8$ which is slightly brighter than the cut-off for our sample.

2.4.3.3 Dependence on spectral type

As a further exploration of the applicability of these methods, we divided our sample into early and late type galaxies based on the `tb` spectral type parameter which describes the SED template fit used by `BPZ` in determining the photometric redshift. Figure 2.8 shows the distribution of spectral type parameters among candidates selected in the BoRG[z8] and BoRG[z9] samples. Over the combined sample this amounted to 390 early type and 585 late type candidates with mean photometric redshifts of $\bar{z}_{early} = 1.89$ and $\bar{z}_{late} = 2.11$ respectively. Considering these samples separately yields galaxy biases of $b_{early} \approx 4.06 \pm 0.67$ and $b_{late} \approx 2.98 \pm 0.98$ when calculated with the R10 method (see table 2.3).

TABLE 2.2: Final values for the galaxy bias combining data from both BoRG[z8] and BoRG[z9] samples with the individual values from the BoRG[z8] sample for reference. The second column shows values derived with the method outlined in LS15, fitting a log-normal distribution to the number counts. The third column contains values derived from the R10 method in which the distribution is assumed to be Poisson plus scatter with the cosmic variance being derived from this excess scatter. Mean photometric redshifts for each sub-sample are given in the final column.

Data set	b_g (LS15)	b_g (R10)	Mean z
Overall	2.94 ± 0.41	3.63 ± 0.57	2.02
BoRG[z8]	2.88 ± 0.60	3.20 ± 0.68	1.99
BoRG[z9]	2.93 ± 0.61	3.94 ± 1.02	2.05

TABLE 2.3: Bias values for early and late spectral type sub-samples of the combined BoRG sample.

Spectral type	b	Mean z
Early	4.06 ± 0.67	1.89
Late	2.98 ± 0.98	2.11

Many previous studies of the clustering of passive and star-forming galaxies at redshift $z \sim 2$ have utilised the well-established BzK photometric selection method (Daddi et al., 2004). Although it is worth noting that this selection yields a slightly different population to that yielded by our selections, clustering studies of passive and star-forming galaxies identified using the BzK method with MUSYC (Blanc et al., 2008) measured a bias of $b_{\rm early} = 3.27^{+0.46}_{-0.47}$ for a sample of passive galaxies at $\bar{z} = 1.58$ and a bias of $b_{\rm late} = 2.93^{+0.59}_{-0.60}$ for a sample of star-forming candidates at $\bar{z} = 1.78$. Notwithstanding the differences in the construction of these galaxy samples, these values are broadly in line with our measurements, demonstrating that the R10 method yields reasonable bias estimates with sufficient pointings.

A clear clustering segregation based on spectral type can be seen in Table 2.3. This segregation whereby early-type galaxies are more strongly clustered than late-type galaxies has long been observed by previous studies in the local Universe (Hermit et al., 1996; Madgwick et al., 2003) and at $z \gtrsim 1$ (e.g. Blanc et al., 2008). These measurements indicate that passively evolving early-type galaxies preferentially inhabit dense regions in large virialised structures. This suggests a link between environment and the quenching of star-formation in galaxies and that interactions that occur in dense environments play an important role in the process of quenching star-formation.

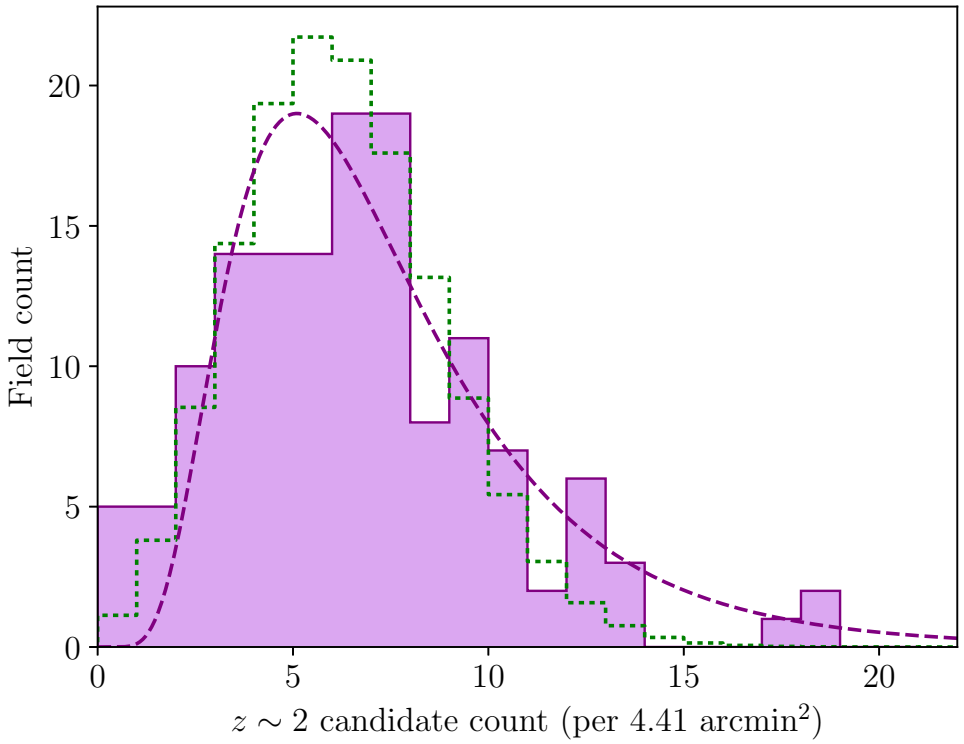

FIGURE 2.6: Distribution of number of counts per field (rescaled to area of 4.41 arcmin^2) after combining counts from the BoRG[z8] and BoRG[z9] samples. Dotted histogram shows Poisson distribution with same mean and normalisation. Excess scatter due to cosmic variance is clearly seen in the distribution of counts. Dashed line shows the best-fit log-normal to the scaled field counts shown.

2.4.4 Comparison of the Two Methods

While the two methods arrived at values that were broadly in line with the values we expect from literature measurements, they did not agree within 1σ with each other for the combined sample. We discuss the reasons behind this disagreement here.

The key difference between the BoRG survey used here and ALHAMBRA survey originally used by LS15 is the volume sampled by each survey. The ALHAMBRA consisted of 8 fields with areas ranging from $720-1500$ arcmin^2 with the survey covering a total of 2.381 deg^2 (8571.6 armin^2). For the bias analysis, their sample was further divided, with each field being divided into sub-fields around $\sim$180 arcmin^2 in area (resulting in a total of 48 sub-fields). This contrasts starkly with the BoRG survey in which the total area sampled is approximately $\sim$700 arcmin^2 with the median field area being 4.41 arcmin^2.

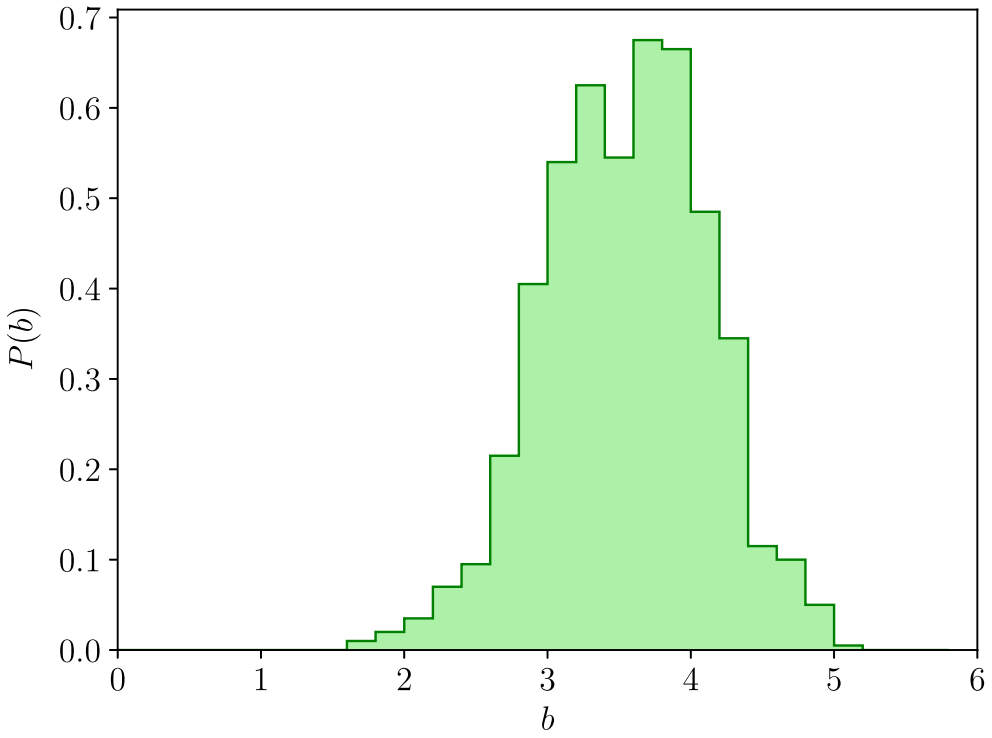

FIGURE 2.7: PDF of galaxy bias values obtained from Monte-Carlo Bootstrapping error analysis for the final value from the combined analysis using the R10 method.

An important consequence of this is that the number counts obtained in our sample are much lower and in particular, the scatter within our sample suffers from a much higher contribution from Poisson shot noise. While we found that the LS15 method still yielded reasonable results for our complete combined sample, when we divided our sample into two sub-samples based on spectral type (section 2.4.3.3), the method appeared to break down ($b_{\rm early} = 1.39 \pm 0.95$ and $b_{\rm late} = 2.57 \pm 0.52$ with the LS15 method). This is presumably due to the fact that 390 early type candidates across 141 BoRG fields (at a mean number count of 2.77 candidates per field) does not provide a sample with sufficiently high number counts in each field to adequately describe the distribution with a log-normal fit. Indeed, the most populated bin for early type candidates is $0 \leq N_{\rm early} < 1$ (using integer bin edges — see figure 2.9). This conclusion is supported by the fact that the LS15 method yielded a much more reasonable bias value for the late-type population where number counts were higher (585 late-type candidates across the sample; only three fields registering no late-type candidates).

The R10 method does not enforce a distribution on the set of number counts, and thus it

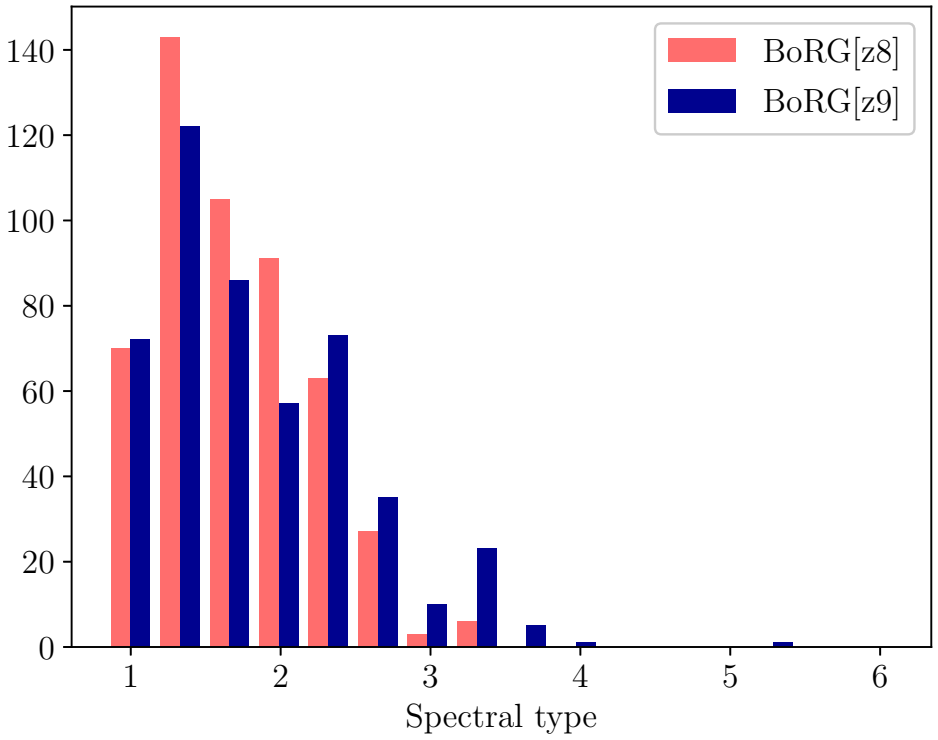

FIGURE 2.8: Distribution of spectral templates used by `BPZ` for $z \sim 2$ candidates in each of the BoRG[z8] and BoRG[z9] samples. Early type galaxies are represented by a template of 1. All templates greater than 1 correspond to late-type and star-forming galaxies with the higher type numbers representative of bluer spectra. `BPZ` discretely interpolates templates, thus fractional type values correspond to sources assigned two templates at a one-to-two weighting. We take candidates with `tb` ≤ 1.33 as early-type and `tb` ≥ 1.67 as late-type. For a full description of these templates, the reader is referred to Coe et al. (2006).

has the advantage of being more robust in samples with low number counts where Poisson shot noise has a more significant impact. Considering we obtained reasonable results with both methods for the full sample, but were only able to obtain reasonable results with the R10 method after dividing the sample based on spectral type, we suggest that our sample lies somewhat near the boundary of what the LS15 method can be applied to in terms of typical galaxy number counts per field.

2.5 Conclusions

In this paper we have presented observational applications of two different methods outlined by López-Sanjuan et al. (2015) and Robertson (2010), in which galaxy bias is measured from the enhanced scatter of a counts-in-cells distribution of galaxies due to

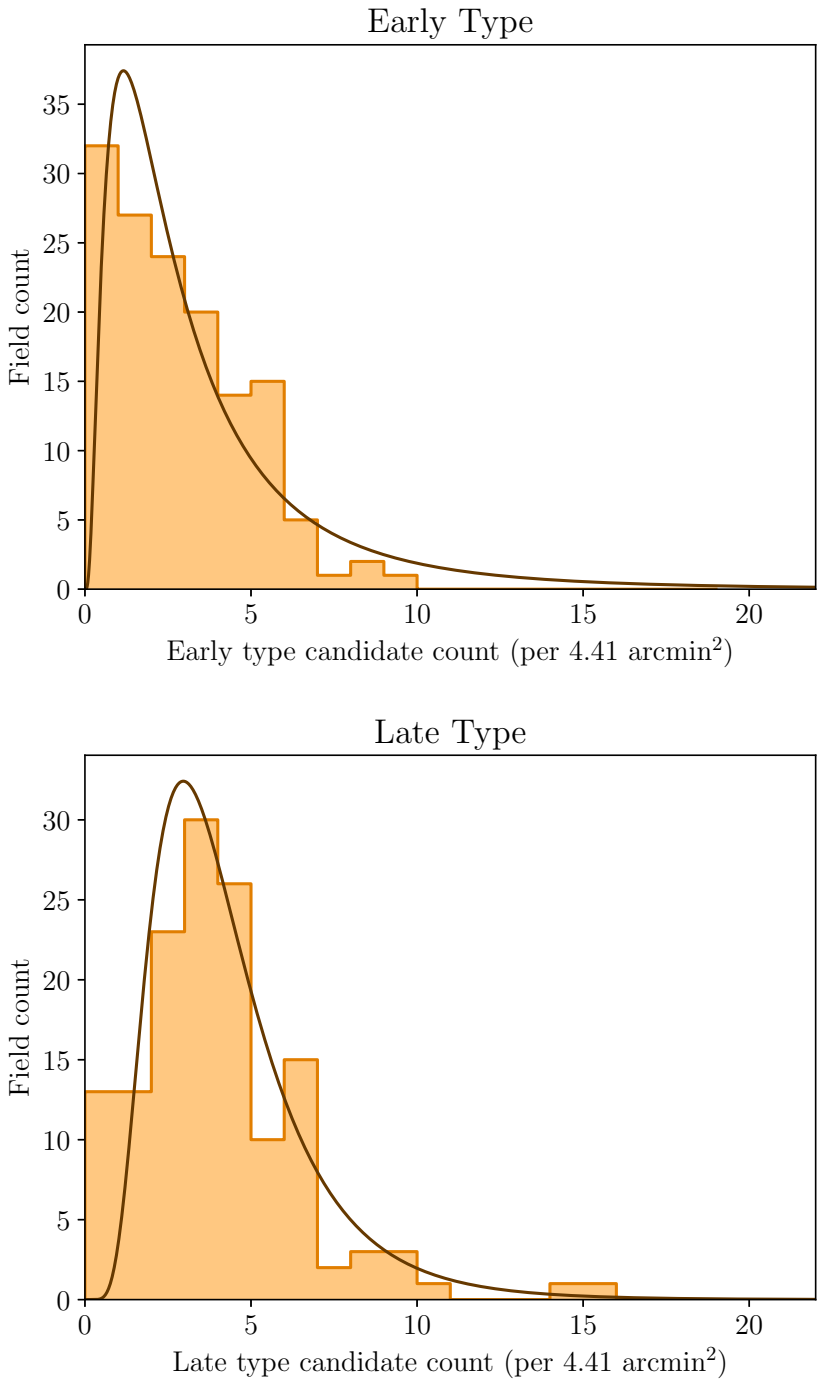

FIGURE 2.9: Histograms showing the distribution of candidate counts per field (scaled to 4.41 arcmin^2) after dividing sample into early- and late-type candidates (based on `tb` parameter from `BPZ` output catalog). The solid curves show the best-fit log-normal distributions to these counts. In the top panel, $0 \leq N_{\text{early}} < 1$ is the most populated bin in the early-type distribution, and thus the log-normal fit presents significant systematic discrepancies that lead to an underestimation of the distribution variance. (Note: analysis is not conducted with the fits shown, these are for illustrative purposes only. Analysis is conducted by fitting a normal to log-densities with the López-Sanjuan et al. (2014) MLE.)

'cosmic' variance. We have applied these methods to photometric candidates at redshift $z \sim 2$ from the *Hubble Space Telescope*'s Brightest of Reionizing Galaxies (BoRG) pure-parallel survey, the first application of these methods to such a survey. The BoRG survey consists of 141 independent fields covering ∼700 arcmin^2 with multiple broad-band optical and near-IR filters in the wavelength range 0.3 – 1.7 μm on WFC3.

Using `SExtractor` we identified resolved sources in the H_{160} near-IR filter and obtained multi-band photometry over this wavelength range. Considering only sources brighter than $m_{AB} = 24.5$, we applied a colour selection ($Y_{098} - H_{160} > 1.5$ for BoRG[z8], $Y_{105} - H_{160} > 1.2$ for BoRG[z9]) to preselect for $z \sim 2$ galaxies with the rest-frame 4000 Å break spectral feature falling between these bands.

We obtain photometric redshifts for sources passing this cut using `BPZ`, identifying $N \sim 1000$ candidates across the 141 fields in the survey with a 'best redshift' in the redshift range $1.5 < z < 2.5$, yielding a mean redshift $\bar{z} = 2.02$ and standard deviation $\sigma_z = 0.25$, after applying quality-of-fit restrictions.

We applied the two methods to the resultant set of field counts, yielding galaxy bias values of $b \approx 2.87 \pm 0.43$ and $b \approx 3.63 \pm 0.57$ for the LS15 and R10 methods respectively for the complete sample, broadly consistent with measurements of the bias at this redshift conducted with correlation functions (*e.g.* Wake et al., 2011). Dividing our sample into $N \sim 400$ early and $N \sim 600$ late type candidates based on SED template fits used in the photometric redshift determinations, we then examined the dependence of the bias on spectral type. This yielded galaxy bias values of $b_{early} \approx 4.06 \pm 0.67$ and $b_{late} \approx 2.98 \pm 0.98$ with the R10 method, in line with correlation function measurements of pBzKs and sBzKs at $z \sim 2$ (*e.g.* Blanc et al., 2008). We found that this divided sample yielded number counts that were simply too low for an adequate log-normal fit to be obtained, preventing the proper implementation of the LS15 method to this part of the analysis.

Comparing the performance of the two methods, we conclude that the LS15 method works best on samples with higher number counts per field, where the effects of Poisson shot noise are suppressed compared to the effects of cosmic variance. The simpler R10 method is more robust for applications where number counts per field are low. We conclude that our sample lies somewhat near the crossover where both methods can reasonably be applied, albeit yielding marginally different results.

The broad agreement of the galaxy bias estimates we have obtained with the LS15 and R10 methods with existing measurements conducted with correlation functions demonstrates observationally for the first time the viability of this method as a tool to measure the bias of a galaxy population at high redshift from cosmic variance with space-based pure-parallel surveys.

This is of particular interest with the upcoming launch of the *James Webb Space Telescope* (*JWST*), which will revolutionise high-redshift galaxy studies with its vastly improved wavelength coverage in the near-infrared. The BoRG survey employed *VYJH* photometry with *HST*/WFC3 which affords NIR coverage to a wavelength of around 1.6 micron, covering rest-frame wavelengths of ∼0.2-0.55 micron for the $z \sim 2$ galaxies presented in this study. This enabled candidates to be selected based on the 4000 Å break. While *HST* lacks the wavelength coverage to extend this technique to targets at higher redshifts, broad band imaging with *JWST*/NIRCam will have coverage to 5 micron, enabling this technique to be applied up to $z \sim 8$. The rest-frame coverage of the *HST*/WFC3 *VYJH* filters for a $z \sim 2$ target would roughly correspond to that of the F200W, F277W, F356W and F444W filters respectively for a $z \sim 8$ target with *JWST*/NIRCam.[2]

Predictions for *JWST* based on the Mason et al. (2015) luminosity function model indicate that a multi-band NIRCam pure-parallel survey similar to BoRG could yield ∼50 $z \sim 6$ and ∼7 $z \sim 8$ candidates per field per 3 hours of observation time to magnitude limit of $m_{\rm AB} \lesssim 28$. It is worth considering that, since the galaxy bias of galaxy populations at different redshifts is likely to be different, evolution of the galaxy bias may affect the way these methods are applied. Qiu et al. (2018) measured the galaxy bias at a range of redshifts from $z \sim 4-7$ segregated by luminosity. The bias measured for their most luminous galaxies (which would be comparable to our selection from the medium-deep BoRG survey) increased from $b = 4.1^{+0.7}_{-0.8}$ at $\bar{z} = 3.8$ to $b = 8.6^{+2.2}_{-2.8}$ for $\bar{z} = 6.8$. Increased galaxy bias would manifest in our field-by-field number count distribution as an increase in the "excess variance" (associated with cosmic variance) that contributes in addition to the expected Poisson variance associated with observing discrete number counts in a random pointing survey (refer to Section 2.3.4.1). However, even if we assume a galaxy bias of $b \approx 8-10$, in line with the highest redshift interval

[2]The precise colour cuts would need to be recalibrated based on specific filter throughputs. Furthermore, systematic differences in SEDs between the $z \sim 8$ and $z \sim 2$ galaxy populations may impact the precise photometric selection criteria.

from Qiu et al. (2018), with mean number counts of $N \sim 7$ we would still expect the number count variance to be dominated by Poisson noise. Accordingly we expect the R10 method would remain robust for application to $z \sim 8$ galaxies observed with *JWST*, provided the number of fields observed was comparable to the BoRG survey. Counts of this order would thus enable the measurement of the galaxy bias with the counts-in-cell approach used in this paper, achieving comparable uncertainties to those reported here. As such, these measurements would be a valuable tool for constraining models of the formation and evolution of galaxies during the Epoch of Reionization.

Chapter 3

Prospects for extending the mass-metallicity relation to low mass at high-redshift: a case study at $z \sim 1$

We report J-band MOSFIRE spectroscopy of a low-mass ($\log(M_*/M_\odot) = 8.62^{+0.10}_{-0.06}$) star-forming galaxy at $z = 0.997$ showing the detection of [N II]and [S II]alongside a strong Hαline. We derive a gas-phase metallicity of $\log(\mathrm{O/H}) = 7.99^{+0.13}_{-0.23}$, placing this object in a region of $M_* - Z$ space that is sparsely populated at this redshift. Furthermore, many existing metallicity measurements in this $M_* - z$ regime are derived from only [N II] /Hα(N2), a diagnostic widely used in high-redshift metallicity studies despite the known strong degeneracy with the ionization parameter and resulting large systematic uncertainty. We demonstrate that even in a regime where [N II]and [S II]are at the detection limit and the measurement uncertainty associated with the [N II] /[S II]ratio is high ($S/N \approx 3$), the more sophisticated Dopita et al. diagnostic provides an improved constraint compared to N2 by reducing the systematic uncertainty due to the ionization parameter. This approach does not, however, dispel uncertainty associated with stochastic or systematic variations in the nitrogen-to-oxygen abundance ratio. While this approach improves upon N2, future progress in extending metallicity studies into this low-mass regime will require larger samples to allow for stochastic variations, as well as careful consideration of the global trends among dwarf galaxies in all physical parameters, not just metallicity.

3.1 Introduction

Measurements of the gas-phase oxygen abundance (metallicity hereafter) in galaxies provide powerful insights into the galaxy-scale star-formation and gas-flow processes that have shaped the development of the galaxy population across cosmic time. Despite extensive studies into the tight correlation between metallicity and stellar mass in galaxies, so-called Mass-Metallicity Relation (MZR), much debate still exists as to its origin (Lequeux et al., 1979; Skillman et al., 1989; Tremonti et al., 2004; Berg et al., 2012; Andrews & Martini, 2013; Yabe et al., 2014; Sanders et al., 2015; Maiolino & Mannucci, 2019). The MZR is observed to have evolved with redshift, with lower average metallicities observed at earlier cosmic times for a given stellar mass (Savaglio et al., 2005; Erb et al., 2006a; Maiolino et al., 2008). However, constraints on the evolution of shape and scatter of the MZR are less clear, due in particular to the difficulties associated with making individual measurements of low-mass galaxies at high redshift. While average metallicity evolution is an important input into galaxy evolution models, extending constraints on chemical evolution to lower mass objects promises key insights into the evolution of the galaxy population.

Existing studies suggest that the shape of the MZR is not constant across cosmic time (Zahid et al., 2013, 2014b). These studies support a downsizing scenario in which low-mass galaxies enrich onto the local MZR at later times. However, despite predictions from theory suggesting that measuring the shape and scatter of the MZR below $\log(M_*/M_\odot) < 9.0$ provide the best prospects for disentangling the driving forces behind this evolution (e.g. Davé et al., 2012), high-redshift studies rarely extend into this mass regime. Thus, the MZR is poorly constrained for low-mass galaxies at high-redshift ($z \gtrsim 1$) as these observations are difficult to carry out, leading to small sample sizes of often loosely constrained measurements (Zahid et al., 2011; Wuyts et al., 2012, 2014, 2016; Belli et al., 2013; Henry et al., 2013a,b; Yuan et al., 2013; Amorín et al., 2014; Maseda et al., 2014). Larger samples of low-mass measurements are required to disentangle the impact of different processes on the evolution of galaxies across cosmic time.

An additional challenge in studies of the MZR below $\log(M_*/M_\odot) < 9.0$ lies with uncertainties in how the metallicities are derived. Diagnostics based on electron temperature (T_e) are widely considered the most reliable measures of metallicity (e.g., review by

Maiolino & Mannucci, 2019). However the range of application of this "direct" method is limited to objects in which the weak [O III] λ4363 emission line can be observed, making it unfeasible beyond moderate redshifts (Jones et al., 2015b; Ly et al., 2016a,b; Calabrò et al., 2017). Motivated by these difficulties, numerous diagnostics have been developed based on ratios of the most easily detected strong rest-frame optical emission lines, calibrated from stellar population synthesis and photoionization models (e.g. Kewley & Dopita, 2002) or T_e measurements taken in either the local universe (e.g. Pettini & Pagel, 2004) or at moderate redshifts (Jones et al., 2015b). Alternatively, a number of generalized approaches exist that compare a range of strong-line fluxes to photoionization models to simultaneously fit for metallicity alongside other key physical parameters (e.g. Pérez-Montero, 2014; Blanc et al., 2015; Vale Asari et al., 2016). As metallicity studies are extended to higher redshifts and lower masses, these strong-line methods become an essential tool in understanding the galaxy population.

Strong-line diagnostics greatly extend the range of stellar masses and redshifts over which metallicities can be derived, however questions remain about their reliability (Kewley & Ellison, 2008; Steidel et al., 2014). In particular, when the number of observed emission lines is small, strong-line diagnostics often fail to disentangle the degeneracy between metallicity and the effects of other physical parameters such as ionization parameter, electron density, hardness of the ionizing sources and relative abundance ratios (e.g. Morales-Luis et al., 2014; Maiolino & Mannucci, 2019). These issues are omnipresent in studies targeting low-mass galaxies at high-redshift where observational challenges frequently limit the range of available emission lines, meaning metallicities are often derived simply from [N II]/Hα (N2; Pettini & Pagel 2004).

The N2 ratio has its advantages in that the two lines are close in wavelength, such that they can be obtained in a single exposure and the ratio is independent of reddening. Thus it has been useful in expanding measurements of the metallicity to faint galaxies (e.g. Erb et al., 2006a; Yabe et al., 2014). However, metallicities derived from this line ratio contain large systematic uncertainties. The [N II] flux is sensitive to the nitrogen abundance and the efficacy of the N2 ratio as a metallicity diagnostic relies on an observed correlation between N/O ratio and metallicity (O/H) (e.g. Pérez-Montero & Contini, 2009; Pérez-Montero, 2014). Any scatter in the N/O vs O/H relation will introduce systematic uncertainty into metallicities derived from the N2 ratio. Additionally, there is a strong

degeneracy between metallicity and ionisation parameter when deriving metallicity from the N2 diagnostic (e.g. Kewley et al., 2019a).

Of greater concern for high-redshift studies is that N2 scaling relations for the star-forming population of galaxies have been shown to evolve with redshift (Kewley et al., 2013; Steidel et al., 2014; Strom et al., 2017), casting doubt on our ability to apply metallicity calibrations derived for the local Universe at high-redshift. Explanations for this offset typically involve possible redshift evolution of some combination of the ionisation parameter (e.g. Bian et al., 2020), N/O to O/H ratio (e.g. Steidel et al., 2014), or Fe/O ratio (which affects the shape of the ionising spectrum; Topping et al. 2020a,b). Accordingly, mitigating against the effects of evolution in these other physical parameters is critical when deriving metallicity from diagnostics that involve the use of the N2 ratio.

In general, measuring larger suites of emission lines will likely be critical to provide robust metallicity measurements, thereby improving constraints on the chemical evolution of galaxies. In line with this, Dopita et al. (2016) (D16 hereafter) have proposed that the set of Hα, [N II] λ6584 and [S II] $\lambda\lambda$6717, 6731 rest-frame optical lines will prove convenient in this pursuit. The relatively narrow wavelength range covered by Hα, [N II] and [S II] lends itself kindly to high-redshift studies as derived line ratios are almost independent of reddening and can typically be observed in one spectroscopic exposure. Additionally, provided the [N II]/[S II] line ratio can be adequately constrained, systematic variation in the derived metallicity caused by degeneracy with ionization parameter and interstellar medium (ISM) pressure is reduced, significantly improving uncertainty as compared to methods utilizing N2 in isolation.

In this contribution, we report J-band spectroscopy taken with *Keck*/MOSFIRE of a low-mass ($\log(M_*/M_\odot) \sim 8.6$) star-forming galaxy at $z \sim 1$ covering the [N II] and [S II] emission lines. With a moderate integration time, we achieve an improved constraint on metallicity using the D16 diagnostic compared to a diagnostic based on N2 alone (once systematic uncertainty is considered). In addition, the combination of ground and space-based spectroscopy covers a suite of line flux measurements that is unique for a $z \sim 1$ dwarf galaxy, allowing us to test consistency of a handful of metallicity diagnostics. Based on this finding, we suggest that targeted surveys utilizing existing

cutting-edge instruments could leverage this diagnostic to place powerful constraints on the processes that govern the evolution of galaxies across cosmic time.

The paper is structured as follows. In Section 3.2 we provide details on the collection of the near-infrared spectral data. Section 3.3 describes our analysis of the data. Section 3.4 presents a brief discussion of the results before we sum up in section 3.5. Throughout this letter we adopt the Planck Collaboration et al. (2016) cosmology: $\Omega_\Lambda = 0.692$, $\Omega_M = 0.308$, $\sigma_8 = 0.815$, and $H_0 = 67.8$ km s^{-1} Mpc^{-1}. All magnitudes are quoted in the *AB* magnitude system (Oke & Gunn, 1983). Unless otherwise stated, [N II] and [S II] refer to [N II] λ6584 and [S II] $\lambda\lambda$6717, 6731 respectively.

3.2 Data

MACS0744_667.0 is a star-forming galaxy at $z = 0.997$ with $m_{\rm AB} = 23.35 \pm 0.02$ in J-band (*HST*/WFC3 F125W) and $R_{\rm eff} = 2.49$ kpc, magnified 1.4× by cluster MACS0744 (lens modeling from Hoag et al. 2019). The source was selected from the HST Grism Lens-Amplified Survey from Space (GLASS; Treu et al., 2015) as an intermediate-redshift target for spectroscopic follow-up with MOSFIRE at Keck (program #Z045M, PI Trenti). Observations were carried out on March 20th, 2016 under good seeing conditions ($\sim$ 0.″4-0.″7 in J band), low atmospheric attenuation ($\Delta m < 0.1$), and minimal airmass (1.05-1.15) for a total of 8457s, divided into individual exposures of 120s each. An ABBA dither pattern with 3″ nodding along the slit was employed and observations started at 19:35HST after acquisition of a standard star during twilight. The MOSFIRE mask included two stars ($m_J = 16.1$ and $m_J = 16.3$) inside the GLASS HST/WFC3 field of view, that were used to verify source alignment.

The MOSFIRE data were reduced using the publicly available data reduction pipeline (DRP[1]). The DRP performs wavelength calibration, rectification, background subtraction and skyline subtraction for each 2D slit in the multi-object slit-mask. The resulting outputs of the DRP are individual 2D signal and noise spectra in electrons per second for each slit on the mask.

[1] https://www2.keck.hawaii.edu/inst/mosfire/drp.html

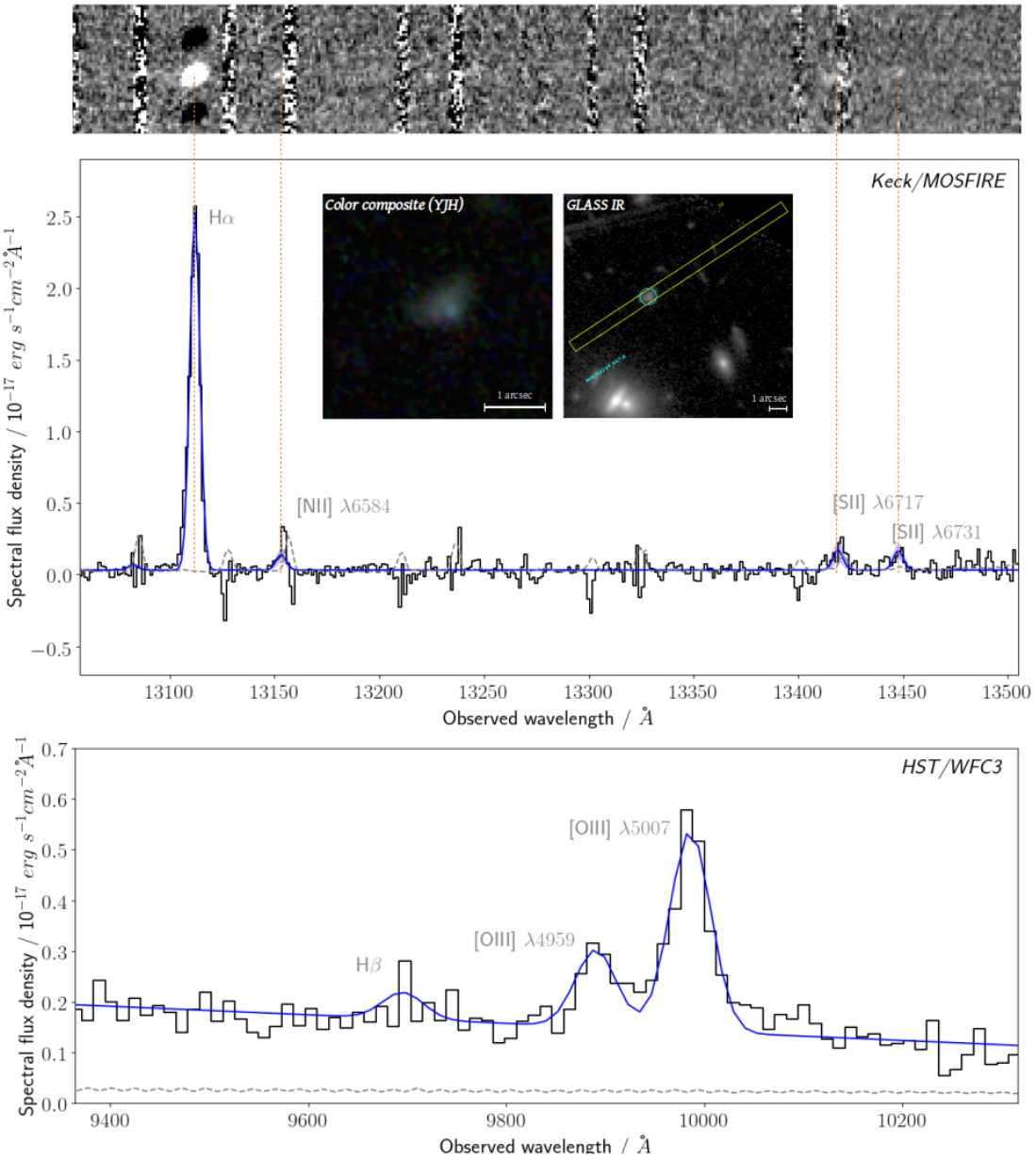

FIGURE 3.1: *Top:* zoom of area of interest in 2D J-band spectrum from MOSFIRE. *Panel 1:* Zoom in on red end of the MOSFIRE J-band spectrum of MACS0744_667.0. Black step-plot shows integrated 1D MOSFIRE signal. Grey dashed line shows 1σ uncertainty associated with each wavelength bin. Blue line shows our best-fit spectrum with 2σ uncertainty depicted by the lighter blue shaded region. Inset is a color composite from CLASH *HST* Y, J, and H band imaging as well as the GLASS alignment reference image showing the MOSFIRE mask slit placement. *Panel 2:* Zoom of lines of interest from integrated 1D *HST*/WFC3 $G102$ grism spectrum of MACS0744_667.0 from the GLASS survey. Color coding is the same as that for panel 1.

3.3 Analysis

3.3.1 1D Spectrum Calibration and Construction

The sensitivity curve for our observations has been derived from a 1D J-band MOSFIRE spectrum of an A5-type reference star (observed with MOSFIRE on 13th June 2016) obtained by integrating over the full width at half maximum (FWHM) spatial extent of the 2D spectrum. We remove the intrinsic stellar spectrum shape by dividing the measured stellar spectrum by an A5-type reference spectrum from the CALSPEC Calibration Database[2]. Thus, the shape of this sensitivity curve gives the response versus wavelength of the set up, due to both the instrument itself and the atmosphere. A faint continuum detection for our object allows us to use its known J-band magnitude (from *HST* imaging) to derive a normalizing factor to calibrate the flux in physical units. Applying this sensitivity curve and normalization factor to the 1D signal extracted by integrating the 2D MOSFIRE spectrum along the 1.12 arcsec spatial extent of the [S II]signal, we recover our 1D J-band spectrum of MACS0744_667.0, shown in Fig 3.1.

3.3.2 Stellar Mass

We derive the stellar mass of MACS0744_667.0 from its spectral energy distribution, following the methods described in our previous analyses of GLASS targets (e.g. Jones et al., 2015a; Hirtenstein et al., 2019). We use 16-band HST photometry spanning observed wavelengths 0.2–1.6 micron from the Cluster Lensing and Supernova Survey with Hubble (CLASH; Postman et al., 2012). The contribution of strong emission lines [O III] , Hβ, Hα+[N II] , and [S II]given in Table 3.1 are subtracted from the broad-band continuum fluxes. Emission line corrected photometry is then fit with the stellar population synthesis code FAST (Kriek et al., 2009). We adopt Bruzual & Charlot (2003) spectral templates with a Chabrier IMF, solar metallicity, Calzetti et al. (2000) dust attenuation curve, and an exponentially declining star formation history. Our analysis is relatively insensitive to adopted stellar metallicity. Assuming a sub-solar $Z = 0.2Z_{\odot}$, comparable to the derived gas-phase value, changes the best-fit stellar mass by only 0.01 dex which

[2] http://www.stsci.edu/hst/observatory/crds/calspec.html

is negligible compared to the uncertainty. Since formal statistical photometric uncertainties typically underestimate the total error (e.g. Ilbert et al., 2006), we scale the uncertainties in flux by a multiplicative factor such that the best-fitting template has a reduced $\chi^2_\nu = 1$. This increases the typical photometric uncertainty from ∼0.03 to 0.08 magnitudes. The resulting best-fit stellar mass is $\log(M_*/M_\odot) = 8.62^{+0.10}_{-0.06}$ after correcting for lensing magnification (MACS0744 cluster lens modelling obtained from Hoag et al. 2019). Note that the subtraction of strong emission lines from the broad-band fluxes reduced the best-fit stellar mass by 0.07 dex.

3.3.3 Line fitting

Our line fitting procedure was run on a wavelength subset of the full J-band spectrum bounded by the bright sky lines at 13055 Å and 13505 Å. Given the continuum in this region is only very tenuously detected, we assume it to be flat at a level taken as the median flux value of the fitting region. After removing this continuum, we obtain line fluxes by fitting a five-peaked Gaussian simultaneously to Hα, [N II]λ6548, [N II]λ6584 and [S II]$\lambda\lambda$6717, 6731 with a χ^2 minimization procedure. To minimize free parameters during fitting, we link all peak centroids and fit only for redshift and assume all peaks have equal line-width. Additionally, the peak height of [N II]λ6548 is assumed to be one third that of [N II]λ6584. Thus we fit for a total of six parameters.

Given the faint [N II]and [S II]emission lines and their proximity to sky lines, we carried out a robust determination and characterization of the errors affecting these flux measurements. In particular, the red shoulder of the [N II]λ6548, [N II]λ6584 and [S II]λ6717 lines were subject to high RMS error according to the MOSFIRE DRP due to possible sky line contamination.

We estimate our uncertainties with a so-called bootstrapping method by perturbing the one-dimensional calibrated spectrum at each wavelength by drawing from a normal distribution with a standard deviation equal to the RMS error assigned to that wavelength by the MOSFIRE DRP. We perform our line-fitting procedure on 1000 realizations of these perturbed 'synthetic' spectra. The mean and standard deviation on the resulting distribution of line-fluxes can be adopted as the measured line flux and its 1σ uncertainty.

We find that when applied to the full fitting range, this bootstrapping method tends to over-estimate the flux by around 2σ when compared to the standard chi-squared fit. However, if we fit again for our set of emission lines, discarding the potentially contaminated values on the red shoulders of the [N II]λ6548, [N II]λ6584 and [S II]λ6717 lines (λ discarded if $\lambda_{centroid} < \lambda < \lambda_{centroid} + 3 \times \sigma_{H\alpha}$; less than 20 values discarded in total), both the bootstrapping and standard chi-squared agree within 1 sigma of the lower value from the original chi-squared fit using all of the values (as opposed to the over-estimated bootstrapping mean value). Thus, we adopt the values and the uncertainty from the bootstrapping method as applied to this amended wavelength set without the discarded points. These line fluxes are given in Table 3.1.

Upon measuring line fluxes, we verify the flux calibration by comparing with measurements from GLASS HST data, noting that the H$\alpha$and [N II]emission lines are blended in the HST spectra. The fluxes are consistent to within $\sim 1.5\sigma$ and the line ratio [S II] /(Hα+[N II]) agrees to within < 0.1 dex.

After subtracting instrument dispersion, estimated at 28.7 km s^{-1} for MOSFIRE J-band, we obtain a rest frame Hαvelocity dispersion of 26.7$\pm$0.5 km s^{-1}. This includes a natural linewidth of $\sigma_0 = 3.2$ km s^{-1} for Hα , and thermal broadening estimated conservatively as $\sigma_{th} = 11 \pm 2$ km s^{-1} (corresponding to nebular temperature $T = 1 - 2 \times 10^4$ K; García-Díaz et al. 2008). Subtracting these effects in quadrature, the intrinsic velocity dispersion of the galaxy is 24.1 ± 0.9 km s^{-1}. Velocity dispersion has been measured for moderate samples of $z \sim 1$ star-forming galaxies by a number of surveys (e.g. KROSS Stott et al. 2016; KLASS Mason et al. 2017). Stott et al. (2016) found a moderate correlation between velocity dispersion and stellar mass in KROSS galaxies at $z \sim 1$. While its low stellar mass places it near the limit of the mass range covered by KROSS, the velocity dispersion measured for MACS0744_667.0 is around a factor of 2 lower than extrapolation of the KROSS $\sigma - M_*$ relation would predict. However, lensed galaxies of similar masses at $z \gtrsim 1$ from the KLASS sample are seen to exhibit comparable velocity dispersion to that found here for MACS0744_667.0 (Mason et al., 2017).

Given the MOSFIRE slit setting (0.″7) and the effective radius of the galaxy (0.″3, derived from HST imaging), we do not expect significant systematic error from the good seeing conditions (∼0.″4-0.″7) during the run as the source image was filling the slit. In addition, the Hαemission line width is resolved at high significance compared to sky lines, and

TABLE 3.1: Fluxes of prominent spectral lines and derived properties of MACS0744_667.0

Spectral line[a]	Flux[b]
Hβ	2.72 ± 0.27
[O III] λ4959	7.93 ± 0.10
[O III] λ5007	20.51 ± 0.80
Hα	14.94 ± 0.11
[N II] λ6584	0.63 ± 0.137
[S II] λ6717	0.84 ± 0.228
[S II] λ6731	0.79 ± 0.161
Derived properties	
$z_{\rm MOSFIRE}$	$0.997 \pm (3 \times 10^{-6})$
$z_{\rm GLASS}$	0.994 ± 0002
$\log(M_*/M_\odot)$	$8.62^{+0.10}_{-0.06}$
$\sigma_{\rm H\alpha}$ / (km s^{-1})	24.1 ± 0.9
$12 + \log({\rm O/H})$	7.99 ± 0.13 [c]
n_e / (cm^{-3})	$\lesssim 1542$ [d]

Notes.
[a] Hβ and [O III] lines obtained with *HST*/WFC3 G102 grism spectroscopy from the GLASS survey. Remaining lines from J-band MOSFIRE spectroscopy.
[b] Fluxes in units of 10^{-17} erg s^{-1} cm^{-2}.
[c] Determined with D16 diagnostic. Quoted uncertainty does not include systematic effects. Refer to §3.3.5 and §3.4 for more information.
[d] Refer to §3.3.6 for details.

the low derived dispersion supports a small dynamical mass. From the half-light radius measured with SourceExtractor (Bertin & Arnouts, 1996b), and following Erb et al. (2006b), we find a dynamical mass of $\log(M_{\rm dyn}/M_\odot) \approx 9.0$. These dynamical results support the SED-based stellar mass derived in §3.3.2, with $\log(M_*/M_{\rm dyn}) \approx -0.4$.

3.3.4 GLASS Line Fluxes

In addition to our MOSFIRE J-band observations, we obtained line fluxes for Hβ and [OIII] from the GLASS slitless spectroscopic observations. The 1D grism spectra from GLASS are included in the high-level science products publicly released by the GLASS team, available from STScI/MAST[3].

The line fitting procedure follows a process similar to that outlined in Section 3.3.3. We fit over the observed wavelength range 9361 Å $\leq \lambda \leq$ 10313 Å, modelling the continuum as a best-fit linear function over this range. We then fit a three-peaked Gaussian profile

[3] https://archive.stsci.edu/prepds/glass/

to the G102 GLASS spectrum (which resolves the [OIII] emission), minimizing free-parameters by fitting for redshift, line-width (assumed equal for all lines), and the areas of each peak.

The low wavelength resolution of the *HST*/WFC3 G102 grism creates difficulties when fitting a continuum, thus uncertainties in the grism line fluxes are likely dominated by uncertainties in the continuum. The uncertainties quoted in table 3.1 are obtained by propagation of the 1σ values obtained for each fit parameter from the co-variance matrix output by the line-fitting function.

In the context of the BPT diagram we find a very high $\log([\mathrm{O\,III}]\lambda5007/\mathrm{H}\beta)$ ratio, perhaps caused by continuum fitting uncertainties. Although the measured position of MACS0744_667 on the BPT diagram is broadly consistent with high ionization $z \sim 2-3$ galaxies observed by Strom et al. (2018).

The flux calibration for the 1D grism spectra from the GLASS data products was carried out independently of this analysis. Comparing emission line fluxes from the GLASS data to our MOSFIRE data we find a very high Balmer decrement ($H\alpha/H\beta = 5.49 \pm 0.55$). We speculate that this measured ratio is unreliable and may arise from differences in the flux calibration or aperture effects caused by comparing the spectrum obatined via slit-spectroscopy with *Keck*/MOSFIRE with slitless data from *HST*/WFC3 G102. Accordingly we pay little heed to emission line ratios derived across these two data sets, instead focusing on ratios derived *within* each data set.

The redshift fit obtained from the GLASS data ($z_{\mathrm{GLASS}} = 0.994$) is slightly offset from that of the MOSFIRE data ($z_{\mathrm{MOSFIRE}} = 0.997$). Given the superior wavelength resolution of MOSFIRE, we take z_{MOSFIRE} to be the source redshift.

3.3.5 Metallicity

The suite of measured line fluxes available to us is quite unique for a galaxy at $z \sim 1$ with $\log(M_*/M_\odot) \leq 9.0$, affording us a range of available metallicity diagnostics. We derive metallicities from diagnostics employing the following line ratios: N2 $= \log([\mathrm{N\,II}]\lambda6584/\mathrm{H}\alpha)$, O3N2 $= \log(([\mathrm{O\,III}]\lambda5007/\mathrm{H}\beta)/([\mathrm{N\,II}]\lambda6584/\mathrm{H}\alpha))$, and N2S2H$\alpha$ $= \log([\mathrm{N\,II}]\lambda6584/[\mathrm{S\,II}]\lambda\lambda6717,6731) + 0.265 \times$ N2. N2 and O3N2 are translated into metallicities using calibrations from Pettini & Pagel (2004) based on a sample of H II regions

TABLE 3.2: Metallicity as derived by different available diagnostics.

Line ratios[a]	Calibration reference	$12 + \log(\mathrm{O/H})$
N2	Pettini & Pagel (2004)	8.11 ± 0.05
O3N2	Pettini & Pagel (2004)	8.01 ± 0.03
N2S2Hα	Dopita et al. (2016)	7.99 ± 0.13

Notes. Quoted uncertainties do not include systematic effects.
[a] Definitions of listed ratio names given in §3.3.5

with direct (T_e) metallicity measurements, while metallicity is inferred from N2S2 using the Dopita et al. (2016) diagnostic, based on theoretical models. These calibrations are as follows:

$$12 + \log(\mathrm{O/H}) = 8.90 + 0.57 \times \mathrm{N2} \tag{3.1}$$

$$12 + \log(\mathrm{O/H}) = 8.73 - 0.32 \times \mathrm{O3N2} \tag{3.2}$$

$$12 + \log(\mathrm{O/H}) = 8.77 + \mathrm{N2S2H}\alpha. \tag{3.3}$$

Applying these diagnostics to our measured line ratios yields values of $Z_{\mathrm{N2}} = 8.11 \pm 0.05$, $Z_{\mathrm{O3N2}} = 8.01 \pm 0.03$ and $Z_{\mathrm{D16}} = 7.99 \pm 0.13$, where $Z = 12 + \log(\mathrm{O/H})$, as given in Table 3.2. Uncertainties quoted here are strictly measurement uncertainties; systematic uncertainties are discussed in §3.4.

3.3.6 Electron Density

The ratio between the [S II]λ6717 and [S II]λ6731 in the [S II]doublet is the most widely used measure of electron density in H II regions. Our detection of this doublet allows us to put constraints on the electron density in this target. We calculate a ratio of [S II]λ6717/[S II]λ6731 $= 1.06 \pm 0.36$. According to the calibration provided by Proxauf et al. (2014a) this places a 1σ upper limit on the density of $n_e \leq 1542$ cm^{-3}.

3.4 Discussion

Strong-line methods are currently the only feasible route to metallicity studies with large samples at high redshifts, particularly for low-mass galaxies. In addition to measurement uncertainty, present at some level in any observation, strong-line measurements in

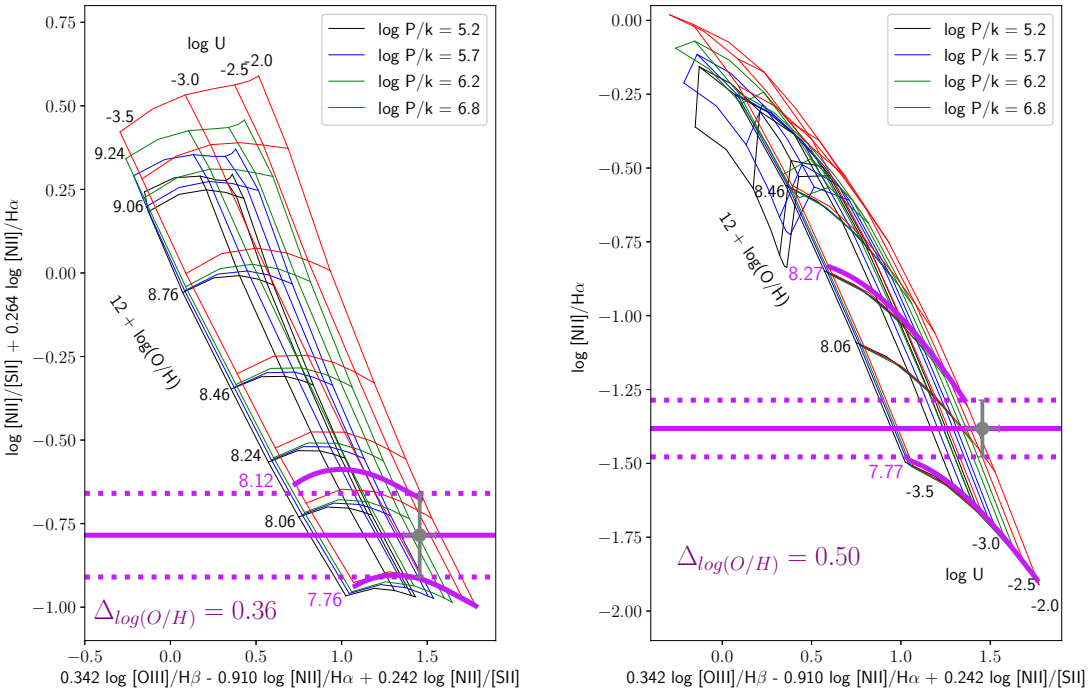

FIGURE 3.2: Two projections of theoretical grids (obtained from Dopita, private communication) showing the improved systematic uncertainty afforded by the D16 diagnostic over the classic N2 diagnostic. The grid lines depict metallicity (roughly horizontal) and ionization parameter (steep diagonal) with different grid colors corresponding to different ISM pressures. The vertical axes in each panel are the set of line ratios required by each respective diagnostic (D16 on the left). The horizontal axis is a combination of line ratios that corresponds neatly with ionization parameter with the D16 y-axis at fixed metallicity and pressure; this is convenient for visual clarity. In both cases the grey point represents measurements reported here for MACS0744_667.0. The horizontal purple lines emphasize the 1σ measurement uncertainty obtained for each quantity given by the vertical axes. These uncertainty intervals clearly indicate the improvement of the D16 diagnostic over the simpler N2 diagnostic. Although the D16 line ratio value itself is not as well constrained (due to lower S/N of the [N II] /[S II]ratio), the reduced systematic uncertainty allows for tighter constraints to be placed on the actual metallicity of this object (1σ upper and lower bounds depicted by purple contours). A first-order approximation with these grids suggests that knowledge of the N2 ratio alone is unable to constrain metallicity tighter than $8.27 \leq 12 + \log(\mathrm{O/H}) \leq 7.77$ while the D16 diagnostic constrains the metallicity to within $8.12 \leq 12 + \log(\mathrm{O/H}) \leq 7.76$ at the 1σ level.

particular suffer from systematic uncertainties caused by degeneracy between metallicity and other physical parameters (ionization parameter, N/O abundance ratio, etc) on the line ratios being employed by the diagnostic. These uncertainties can arise from both stochastic variations of these physical parameters, as well as any systematic variations that may be present in the high-redshift universe. Understanding and minimizing the uncertainties associated with these methods is therefore a critical and open issue.

As a result of the observational challenges associated with assembling large suites of

emission lines for low-mass galaxies, the [N II] /Hαdiagnostic (N2; Pettini & Pagel 2004) is widely used in high redshift studies as the required lines are relatively strong and are close in wavelength (e.g. Erb et al., 2006a; Wuyts et al., 2012; Yabe et al., 2014). By contrast, many widely used strong-line diagnostics, such as O3N2 (Pettini & Pagel, 2004) or R_{23} (Zaritsky et al., 1994), can deliver lower systematic uncertainty by better accounting for degeneracy with other physical parameters, but are considerably more challenging to obtain at high redshifts due to the wavelength coverage required. However, the reliance of N2 on a single line ratio makes it vulnerable to systematic uncertainties; in particular those caused by degeneracy with ionization parameter and N/O abundance ratio.

3.4.1 Ionization Parameter Dependence

Strom et al. (2018) show that ionization parameter is the physical parameter to which nebular spectra respond most sensitively. Accordingly, the dependence of the N2 ratio on ionization parameter introduces systematic uncertainty into metallicities derived from N2. The detection of the [S II]doublet in this object allows us to adopt the D16 approach of incorporating [N II] /[S II]in addition to N2, which drastically reduces degeneracy between metallicity, ionization parameter and ISM pressure.

Grids obtained from M. Dopita (private communication) plotted in Figure 3.2 clearly highlight the short-comings of the classic N2 diagnostic. The grids show how the expected line ratios vary with ionization parameter and ISM pressure at fixed metallicity. We plot MACS0744_667.0 as reported in this letter onto these grids to illustrate the tighter constraint afforded by using the D16 diagnostic. Although the line ratio itself has a larger uncertainty, the vastly reduced ionization dependence leads to an overall better metallicity measurement (constrained to a range of 0.36 dex for D16 c.f. 0.50 for N2; 1σ upper and lower bounds indicated by the purple curves in Fig 3.2). Furthermore, unlike systematic uncertainty, measurement uncertainty reduces with higher signal-to-noise, meaning prospects for placing tight constraints on the metallicities of low-mass galaxies at high-redshift are greatly improved for D16 compared to when the N2 line ratio is used in isolation. This approach is made feasible in this low-mass and high-redshift regime by virtue of the spectral proximity of the [S II]doublet to the H$\alpha$and [N II]λ6584

lines, meaning they can be obtained without having to essentially double time-on-target requirements by observing with additional filters.

Figure 3.3 shows metallicity, including estimated systematic uncertainty due to ionization parameter (dotted error bars), plotted against the stellar mass (see § 3.3.2) for this object and a few available in the literature in a similar redshift range ($z \sim 1$). The black triangles in Fig 3.3 show the $z \sim 0$ MZR derived from the SAMI survey (Sánchez et al., 2019). Given that chemical evolution of low-mass galaxies is expected to be more significant at later times (e.g. Henry et al., 2013b), samples of objects with stellar masses below $\log(M_*/M_\odot) < 9.0$ beyond redshift $z \gtrsim 1$ promise to provide valuable insight into evolutionary processes driving the galaxy population. In Fig 3.3 the limitations caused by systematic uncertainties can be seen. Despite the low measurement uncertainty for our N2 metallicity (left panel), the large systematic uncertainty means the resultant measurement provides little insight to distinguish between cases where significant or very little evolution occurs from $z \sim 1$ to $z \sim 0$.

3.4.2 Nitrogen-to-Oxygen Ratio Dependence

A remaining concern, however, is that the N2 ratio is primarily sensitive to the nitrogen abundance. Thus, oxygen abundances can only be inferred using some (implicitly or explicitly) assumed N/O ratio. Systematic uncertainty on measurements of oxygen abundances conducted in this way can therefore be introduced in two main ways. First, even assuming an appropriate N/H-to-O/H conversion can be applied, stochastic variations of log(N/O) at fixed log(O/H) add to the overall uncertainty of the final metallicity measurement. In a sample of objects with log(N/O) and 12+log(O/H) direct measurements (Berg et al., 2012; Pilyugin et al., 2012) we found that among the objects with $7.8 < 12+\log(O/H) < 8.2$, the standard deviation in log(N/O) was $\sigma_{\log(\mathrm{N/O})} = 0.13$ dex, comparable to line ratio measurement uncertainties. Although strong-line measurements at $z \sim 2.3$ by Strom et al. (2018) suggest this scatter could be as large as 0.8 dex in high-redshift galaxies. This stochastic variation limits the accuracy of individual metallicity measurements; larger samples are required to account for this effect. Second, some authors suggest that the N/O abundance ratio has undergone evolution with redshift (e.g. Masters et al., 2014). Although, Steidel et al. (2016) suggest that $z \sim 2$ galaxies on average lie within the same trend as local galaxies. This will critically affect metallicities

derived using nitrogen lines. However, it is a difficult issue to address, requiring large samples of high quality spectra at high redshift. If, indeed, high-redshift galaxies do exhibit different N/O ratios to local galaxies the assumed N/O parameterization may be inappropriate. Modest samples of "direct" T_e measurements at $z > 1$ may help to further understand this.

3.4.3 Insights into the low-mass end of the mass-metallicity relation

This single object study of MACS0744_667.0 has focused on characterising the overall level of uncertainty on strong-line metallicity measurements that make use of the [N II] emission line, with a particular focus on the systematic uncertainty caused by degeneracy with ionisation parameter. Of course, the goal of metallicity studies is to use oxygen abundance measurements to place constraints on the evolutionary history of galaxies. Naturally, with only the single object presented here, we are unable to place strong constraints on the low-mass behaviour of the $z \sim 1$ mass-metallicity relation. However, in this section we explore the prospective insights that could be obtained from a modest sample (∼10) of similar objects.

Figure 3.3 shows MACS0744_667.0 in the context of the mass-metallicity relation as derived from $z \sim 0$ galaxies from SAMI (black triangles; Sánchez et al. 2019). Where available, a selection of literature measurements at $z \sim 1 - 2$ are also shown for comparison (Wuyts et al., 2012, 2016; Yabe et al., 2014). Existing studies at higher masses at this redshift had already shown an evolution of the mass-metallicity relation toward lower normalisation from $z \sim 0$ to $z \sim 1.5$ (e.g. Yabe et al. 2014; see left panel of Figure 3.3). However, previous studies have lacked the low-mass coverage to adequately constrain the slope evolution of the mass-metallicity relation from $z \sim 0$ to $z \sim 1.5$.

In the so-called "downsizing" scenario of galaxy evolution, the chemical evolution of low-mass galaxies is expected to be more significant at later cosmic times. Indeed, a number of authors have interpreted the evolution of the mass-metallicity relation through this lens, finding steeper slopes at earlier cosmic times (Maiolino et al., 2008; Zahid et al., 2013). Furthermore, equilibrium models from Davé et al. (2012) have been used to distinguish between different feedback prescriptions dominated by energy-driven or feedback-driven winds based on the precise low-mass slope of the mass-metallicity relation at $z \sim 1$ (Henry et al., 2013a,b).

We can see in the left-hand panel of Figure 3.3 that the large systematic uncertainty on the N2-based metallicity for MACS0744_667.0 makes it challenging to distinguish between the scenario of very little slope evolution from $z \sim 0$ to $z \sim 1.5$, and the scenario of significant steepening over that same interval. From this, we conclude that there is very limited utility in attempting to further constrain the shape evolution of the mass-metallicity relation with metallicity measurements based solely on the N2 ratio.

The D16 diagnostic is relatively new and has not been extensively applied. Thus the right hand panel of Figure 3.3 lacks a high-mass $z \sim 1-2$ sample with which to compare our D16 metallicity measurement for MACS0744_667.0. However, with the greatly reduced dependence on ionisation parameter, we can see that the metallicity uncertainty is dominated by measurement uncertainty rather than systematic uncertainty. Thus, prospects for constraining the shape evolution of the mass-metallicity relation with this diagnostic are far more promising.

This study has shown that the [S II] $\lambda\lambda$6717, 6731 emission lines required for the D16 diagnostic can be observed alongside Hα and [N II] λ6584 with moderately deep spectra of gravitationally lensed $z \sim 1-2$ galaxies using modern ground-based near-infrared spectrographs on 8-10 m class telescopes. Furthermore, we have demonstrated here that utilising the additional [N II]/[S II] line ratio leads to a significant decrease in the systematic uncertainty of the derived metallicity. A modest sample ($\sim$10 – 15 galaxies) could thus reasonably be obtained with a multi-object spectrograph on an 8-10 m class telescope, such as *Keck*/MOSFIRE or *VLT*/KMOS. Combined with a reanalysis of a higher mass sample with the D16 diagnostic, such a sample could place improved constraints on the shape evolution of the mass-metallicity relation. This in turn could be used to improve constraints on the feedback mechanisms shaping the evolution of galaxies across cosmic time.

3.4.4 Future Prospects

While better constraints on the Mass-Metallicity relation below $\log(M_*/M_\odot) \leq 9.0$ at $z \gtrsim 1$ promise unique insights into evolution of the galaxy population, progress has been limited by the associated observational challenges. Strong-line metallicity measurements are the only feasible approach to making progress in this area, however the systematic

uncertainties associated with the simple N2 diagnostic clearly limit its effectiveness in distinguishing between different evolutionary processes.

An additional source of uncertainty not discussed here is the contribution to the measured [S II] /Hα ratio from diffuse ionized gas (DIG). The consequence of this is that global metallicities derived from N2S2Hα will be sensitive to variations in the fraction of DIG (f_{DIG}) in the galaxy population. As highlighted in Shapley et al. (2019), if high-redshift galaxies follow the same relation between f_{DIG} and Σ_{SFR} as local galaxies, N2S2Hα would vary systematically with redshift and thus not be appropriate for comparing high- and low-redshift samples. Thus, further observations are required to determine the degree to which variations in f_{DIG} would affect the systematic uncertainties induced in a sample at fixed redshift beyond $z \gtrsim 1$.

Indeed, more generally, it is important to note that while including the [N II]/[S II] ratio may decrease the impact of systematic uncertainty, it certainly does not outright remove its presence. We can consider for example differences between the absolute metallicities for massive galaxies derived for SAMI galaxies from the N2 and D16 diagnostics (left and right hand panels of Figure 3.3 respectively; Sánchez et al. 2019). For galaxies in that sample with masses greater than $\log(M_*/M_\odot) \gtrsim 10.5$, we can see that metallicities derived from D16 are higher by a factor of $\sim$2 than those derived from N2. If we consider as an example a Milky Way mass galaxy ($\log(M_*/M_\odot) = 10.78$; Licquia & Newman 2015), we could expect typical metallicities from this sample of $12+\log(O/H) \approx 8.78$ (slightly super solar) when derived from N2. Using the D16 diagnostic instead predicts a higher value of $12+\log(O/H) \approx 9.15$ (super solar by a factor of $\sim$2). If we assume the metallicity of the Milky Way is approximately solar and that it represents a typical galaxy with $\log(M_*/M_\odot) = 10.78$, it would suggest that the N2 value is a better estimate of the absolute metallicity in this high metallicity regime. However, we note that the normalisation of mass-metallicity relations derived from strong-line methods have long been known to be highly dependent on diagnostic (Kewley & Ellison, 2008). Accordingly some authors have suggested that interpretation of strong-line method based mass-metallicity relation studies is best limited to evaluation of *relative* metallicities within large samples, rather than absolute metallicities of individual objects (e.g. Sanders et al., 2020b). Nonetheless, we should be mindful that the D16 diagnostic is by no means free of systematic uncertainty. More high quality data would be needed at high-redshift to better characterise its performance in studies such as this one.

Approaches that better constrain degeneracy by including more emission line ratios are certainly preferable, although challenging in this regime where long integration times are required to accurately measure even the strongest metal emission lines. Approaches that use photoionization models to simultaneously fit for all of these physical parameters including metallicity (e.g. Pérez-Montero, 2014; Blanc et al., 2015) appear to lend themselves naturally to this context, however we did not include these in this discussion as the low measured S/N for MACS0744_667.0 implies that both the specific details of the input model adopted, and parameter priors could have a substantial impact on the inference. While future facilities will certainly aid progress in this area, applying the D16 diagnostic to deep observations with existing instruments can improve the ionization dependence of existing N2-based constraints without requiring the factor of 2-3 increase in time-on-target associated with many other strong-line methods.

3.5 Conclusion

Extending constraints on the high-redshift ($z \gtrsim 1$) Mass-Metallicity Relation to masses below $\log(M_*/M_\odot) < 9.0$ promises powerful insight into the evolutionary processes that govern the galaxy population. Currently strong-line methods are the only viable approach for expanding metallicity measurements to high-redshift dwarf galaxies.

However, particularly when the number of available emission lines is small, metallicity measurements made with strong-line methods may suffer from degeneracy with other physical parameters such as ionization parameter, chemical abundance ratios and ISM pressure. Derived metallicities can be affected by either stochastic or systematic variations in these properties among the high-redshift galaxy population.

In this contribution we have presented MOSFIRE J-band spectroscopy of MACS0744_667.0, a low mass ($\log(M_*/M_\odot) = 8.62^{+0.10}_{-0.06}$) star-forming galaxy at redshift $z = 0.997$ magnified 1.4× by CLASH cluster MACS0744 in which we observe detection of [N II]λ6584 and [S II]$\lambda\lambda$6717, 6731 alongside strong Hαdetection. Additionally, we derive Hβ, [O III]λ4959 and [O III]λ5007 line fluxes from HST/WFC3 G102 grism spectroscopy from the GLASS data release. Access to this set of emission lines is quite unique for a galaxy of this mass at this redshift.

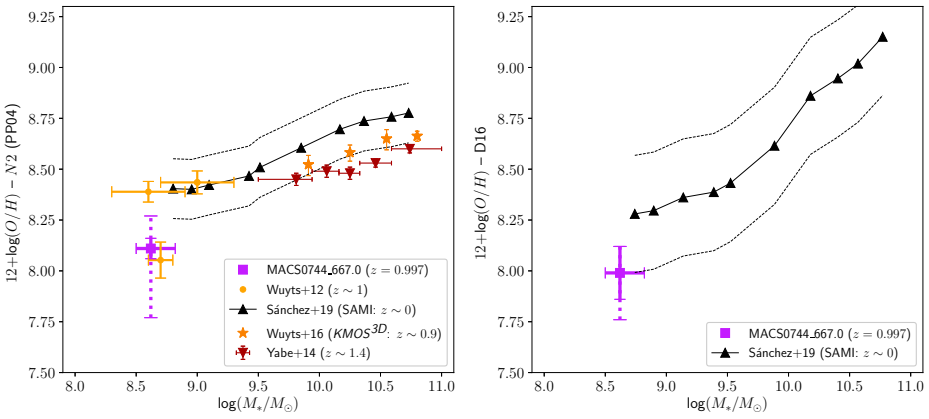

FIGURE 3.3: MACS0744_667.0, shown as the purple square, resides in a region of $M_* - \log(O/H)$ space that is sparsely populated by existing observations. Solid error bars show 1σ measurement uncertainty while the dotted error bars depict our estimate of additional systematic uncertainty due to uncertainty in the ionization parameter (see Fig 3.2). *Left panel:* Mass-metallicity relation with metallicities derived from the N2 diagnostic (Pettini & Pagel, 2004). Orange stars are metallicities derived from N2 ratios of stacks of spectra at $0.6 < z < 1.1$ binned by mass from $KMOS^{3D}$ (Wuyts et al., 2016). Orange circles are N2 metallicities from individual lensed objects at $z \sim 1$ from Wuyts et al. (2012). Red inverted triangles are N2 metallicities of $z \sim 1.4$ galaxies binned by mass from Yabe et al. (2014). Displayed error bars reflect only measurement uncertainty for these objects. Black triangles show the N2 gas-phase metallicities binned by mass measured at 1 R_e in SAMI galaxies at $z \sim 0$ with dashed black lines indicating the average residual 1σ scatter after fitting a Mass-Metallicity Relation (Sánchez et al., 2019). Some evolution is seen from $z \sim 1.4$ to $z \sim 0$ above $\log(M_*/M_\odot) \geq 9.5$, however neither large sample at high redshift is able to probe below $\log(M_*/M_\odot) < 9.5$ where evolution is expected to be most significant. Despite the low measurement uncertainty associated with MACS0744_667.0 the large systematic uncertainty limits the degree to which insight can be gained from this measurement. *Right panel:* Mass-metallicity relation with metallicities derived from the D16 diagnostic. Black triangles show the D16 gas-phase metallicities measured at 1 R_e in SAMI galaxies at $z \sim 0$ (Sánchez et al., 2019). Here, MACS0744_667.0 has a larger measurement uncertainty, however the additional systematic uncertainty is less problematic, suggesting larger samples using this diagnostic will be more effective in constraining redshift evolution of the mass-metallicity relation.

We derive metallicity from N2 ($12 + \log(\mathrm{O/H}) = 8.11 \pm 0.05$ with statistical uncertainty; $12 + \log(\mathrm{O/H}) = 8.11^{+0.16}_{-0.34}$ including additional systematic uncertainty) as well as N2S2Hα ($12+\log(\mathrm{O/H}) = 7.99 \pm 0.13$; $12+\log(\mathrm{O/H}) = 7.99^{+0.13}_{-0.23}$). While the inclusion of the [N II] /[S II]ratio in N2S2Hα increases the measurement uncertainty, we find that even in this case where $\mathrm{S/N}_{N2S2}$ is small, the improved mitigation of the dependency on ionization parameter reduces the overall uncertainty on the metallicity measurement. We estimate that uncertainty due to N/O abundance ratio is likely of comparable order to the measurement uncertainty of N2S2Hα. Large samples of high quality spectra of high-redshift dwarf galaxies are needed to assess if there is a systematic variation of

this abundance ratio at high-redshift and at what level the stochastic variation impacts dwarf galaxy metallicities.

Further progress in extending the high-redshift Mass-Metallicity Relation to dwarf galaxies requires careful consideration of the global trends among dwarf galaxies in all physical parameters, not just metallicity. Deeper surveys targeting low-mass objects at $z \gtrsim 1$ employing existing multiplexed NIR instruments (e.g. *Keck*/MOSFIRE or *VLT*/KMOS) will improve understanding of the stochastic variations in these properties among the high-redshift dwarf population, providing unique insight into the evolutionary processes that govern the galaxy population.

Chapter 4

Spatially resolved direct method metallicity in a high-redshift analogue local galaxy: temperature structure impact on metallicity gradients

We investigate how H II region temperature structure assumptions affect "direct-method" spatially-resolved metallicity observations using multispecies auroral lines in a galaxy from the SAMI Galaxy Survey. SAMI609396B, at redshift $z = 0.018$, is a low-mass galaxy in a minor merger with intense star formation, analogous to conditions at high redshifts. We use three methods to derive direct metallicities and compare with strong-line diagnostics. The spatial metallicity trends show significant differences among the three direct methods. Our first method is based on the commonly used electron temperature T_e([O III]) from the [O III] λ4363 auroral line and a traditional T_e([O II]) – T_e([O III]) calibration. The second method applies a recent empirical correction to the O^+ abundance from the [O III]/[O II] strong-line ratio. The third method infers the T_e([O II]) from the [S II] $\lambda\lambda$4069,76 auroral lines. The first method favours a positive metallicity gradient along SAMI609396B, whereas the second and third methods yield flattened gradients. Strong-line diagnostics produce mostly flat gradients, albeit with unquantified contamination from shocked regions. We conclude that overlooked assumptions about the internal temperature structure of H II regions in the direct method can lead to large discrepancies in metallicity gradient studies. Our detailed analysis of SAMI609396B underlines that high-accuracy metallicity gradient measurements require a wide array

of emission lines and improved spatial resolutions in order to properly constrain excitation sources, physical conditions, and temperature structures of the emitting gas. Integral-field spectroscopic studies with future facilities such as *JWST*/NIRSpec and ground-based ELTs will be crucial in minimising systematic effects on measured gradients in distant galaxies.

4.1 Introduction

Abundances of heavy elements (metallicities) in the interstellar medium (ISM) of galaxies are enriched by stellar nucleosynthesis and trace star formation histories and gas-flow processes that ultimately shape the galaxy population. In particular, the spatial distribution of metallicity offers a powerful probe on the role of mergers, outflows, gas mixing, and gas accretion in transforming galaxies (e.g., Edmunds & Greenhow, 1995; Kewley et al., 2010; Torrey et al., 2012; Magrini et al., 2016; Finlator, 2017; Ma et al., 2017; Bresolin, 2019; Tissera et al., 2019; Hemler et al., 2020). Spatial distributions of metals are often summarised as radial abundance gradients and azimuthal variations (e.g., Searle, 1971; Vila-Costas & Edmunds, 1992; Li et al., 2013; Ho et al., 2015, 2019), with both negative and flat metallicity gradients widely observed in the Milky Way and other local galaxies (e.g., Deharveng et al., 2000; Bresolin et al., 2004; Berg et al., 2013, 2020).

Spatially resolved studies of galaxies are now far more accessible compared to a decade ago, thanks to the advent of integral-field unit (IFU) spectroscopy. Multiplexed IFU surveys (e.g. CALIFA Sánchez et al. 2012; SAMI Bryant et al. 2015; MaNGA Bundy et al. 2015) have afforded large samples of gradient measurements in the local Universe. Studies find a dependence on stellar mass: low-mass galaxies ($\sim 10^9$ $M_\odot$) show almost flat gradients, with negative gradients steepening to high masses (Belfiore et al., 2017; Poetrodjojo et al., 2018). Spatially resolved measurements become more challenging at high redshift and observations show a substantial amount of scatter (Yuan et al., 2011; Jones et al., 2013; Leethochawalit et al., 2016; Carton et al., 2018; Wang et al., 2020; Curti et al., 2020b). One major caveat in using this broad range of observations to develop a coherent model of galaxy evolution is that different measurement techniques are often in disagreement.

A number of observational methods exist for determining the oxygen abundance (metallicity hereafter) of the ISM in galaxies from emission line spectroscopy (see Maiolino & Mannucci 2019; Kewley et al. 2019a for recent reviews). However, different techniques often show large offsets up to 0.7 dex (e.g. Kewley & Ellison, 2008; Peimbert et al., 2017). This stark disagreement between different metallicity measurement techniques presents an ongoing challenge for studying chemical evolution of galaxies.

Emission line strengths in the photoionised nebulae around hot O- and B-type stars (H II regions) are sensitive to electron temperature (T_e), in addition to ionic abundances, ionisation parameter, and ISM pressure. Thus, a desirable approach to metallicity measurement is to use ratios of auroral emission lines and corresponding strong nebular emission lines to explicitly determine T_e, and subsequently metallicity (Direct Method; e.g., see Pérez-Montero 2017 for an overview). This "direct method" is traditionally considered the gold standard in abundance determination (e.g. Maiolino & Mannucci, 2019), and underpins the calibration of many alternative techniques (e.g. Pettini & Pagel, 2004; Curti et al., 2020a). However, one major practical issue with the direct method is that the faintness of the optical auroral lines severely limits its application. An alternative T_e-based method outlined by Jones et al. (2020) determines oxygen abundance based instead on far-infrared oxygen lines ([O III] 52μm or [O III] 88μm). This is expected to be favourable beyond $z \gtrsim 5$ where these far-IR features can be observed with millimeter instruments such as *ALMA*, but is difficult to apply at lower redshifts.

Due to the faintness of auroral lines required for the direct method, strong-line methods are widely adopted in observations. Strong-line methods use ratios of the brightest rest-frame ultra-violet and optical emission lines to empirically determine the metallicity with calibrations based on either direct-method observations (e.g. Pettini & Pagel, 2004; Pilyugin & Thuan, 2005; Curti et al., 2020a) or stellar population synthesis and photoionisation models (e.g. Kewley & Dopita, 2002; Kobulnicky & Kewley, 2004; Dopita et al., 2016). Strong-line methods vastly expand the redshift and mass range of galaxies for which metallicities can be derived. However, it has been widely observed that metallicities measured with different methods often disagree (e.g. Kewley & Ellison, 2008; Moustakas et al., 2010; Morales-Luis et al., 2014). In particular, theoretical methods, are reliant on simple geometries, such as spherical or plane parallel, and assume a constant temperature, constant density, or a constant pressure.

Despite the baseline role of the direct method, it does have limitations beyond practical detection-rate issues (Nicholls et al., 2020; Yates et al., 2020). H II regions are complex structures and summarising their conditions with integrated measurements of emission line ratios carries many assumptions. For example, H II regions are known to have internal temperature variations (Peimbert, 1967; Kewley et al., 2019a). An observed emission line ratio samples the luminosity-weighted average conditions of the emitting nebulae (Nicholls et al., 2020). The direct method is best applied by constructing a multi-zone temperature model using auroral lines from multiple ionic species (e.g. Pérez-Montero, 2017; Berg et al., 2020). Commonly used auroral lines include those from O^{2+}, O^{+}, N^{+} or S^{2+} ions.[1] This allows internal temperature gradients to be sampled since ions with differing ionisation energies preferentially sample different sub-regions of the nebulae.

However, measuring auroral lines from multiple species in observations presents a difficult practical challenge. Even detection of a single auroral line, commonly [O III] λ4363, is generally considered a favourable outcome. But since the [O III] λ4363 line is only produced in the hottest regions of a nebula, a resulting T_e-derived metallicity may be a lower limit to the true metallicity if there is a temperature gradient (Kewley et al., 2019a). To overcome the lack of direct constraints on the multi-zone temperature structure, abundance measurements are often made adopting empirical relations between temperatures from different ions. For example, the [O II] temperature (T_e([O II])) is indirectly inferred from the [O III] temperature (T_e([O III]); based on [O III] λ4363) using the T_e([O II]) – T_e([O III]) relation (e.g. Izotov et al. 2006; López-Sánchez et al. 2012; Pérez-Montero 2017). Recently, Yates et al. (2020) show that at low O^{2+}/O^{+}, this approach can lead to large deficits in the measured O^{+} abundance, causing total oxygen abundances to be underestimated by up to $\sim$0.6 dex.

Studying metallicity in spatially resolved detail exacerbates the practical limitations of the direct method. Indeed, direct method metallicities have been mapped only for the Milky Way and small samples of large nearby spiral galaxies (Deharveng et al., 2000; Bresolin et al., 2004; Li et al., 2013; Berg et al., 2013, 2015, 2020; Croxall et al., 2015, 2016; Ho et al., 2019), exploring only a very narrow subset of the galaxy population. Here we leverage public release IFU data from the SAMI Galaxy Survey to expand spatially resolved direct method metallicity measurements to a new parameter space.

[1] The Cl^{2+} and Ar^{3+} ions can provide similar temperature probes to complement O^{2+} measurements, however are usually too faint to be detectable.

From a search of auroral lines in SAMI Data Release 2 data cubes, we identify one particularly strong candidate: SAMI609396. This target is a minor-merger system and one galaxy in the system (SAMI609396B) is experiencing a burst of star formation. SAMI609396B is analogous to a high-redshift galaxy given its low-mass and high SFR. We detect prominent, spatially resolvable emission of three auroral lines: [S II] $\lambda\lambda$4069, 76, [O III] λ4363 and [S III] λ6312 in SAMI609396B.

In this contribution, we focus on this notable case to study direct method metallicity and electron temperature in a spatially resolved manner. The presence of auroral lines from multiple ionic species allows us to investigate the common assumption of using an assumed temperature relation (e.g. T_e([O II]) – T_e([O III]) relation) on the spatial distribution of metallicity in galaxies. Additionally, comparisons to strong-line metallicity trends provide further insight into possible systematic effects in samples of gradients measured in the local and high-redshift Universe. Given the rarity of spatially resolved T_e studies at low redshift, and the relevance of this object to high-redshift comparisons, it warrants a detailed study of its own.

This work is organised as follows. In Section 4.2 we briefly describe the SAMI DR2 public release data, general properties of the SAMI609396 system, and selection of SAMI609396B. Our methodology for deriving spatially resolved electron temperature measurments is outlined in Section 4.3. In Section 4.4 we derive metallicity maps from three different "direct method" approaches and four different strong-line methods and discuss the differences in spatial trends favoured by each. We discuss further caveats in Section 4.5 before summarising and presenting conclusions in Section 4.6. Detailed descriptions of the derivation of global properties, spectral fitting and emission line measurements are deferred to the Appendix. We also include a list of SAMI galaxies with visually identifiable auroral line emission in the Appendix. Throughout this paper we adopt the Planck Collaboration et al. (2016) cosmology: $\Omega_\Lambda = 0.692$, $\Omega_M = 0.308$, $\sigma_8 = 0.815$, and $H_0 = 67.8$ km s^{-1} Mpc^{-1}. All magnitudes are quoted in the *AB* magnitude system (Oke & Gunn, 1983).

4.2 The SAMI Galaxy Survey

We conducted a search for auroral lines in SAMI Galaxy Survey Public Data Release 2[2] (Bryant et al., 2015; Green et al., 2018; Scott et al., 2018). The SAMI Galaxy Survey (Bryant et al., 2015) is a large IFU survey targeting low-redshift ($z \lesssim 0.1$) galaxies with the Sydney – Australian Astronomical Observatory Multi-Object Integral Field Spectrograph (Croom et al., 2012). Reduced SAMI data cubes are formed by sampling dithered hexabundle observations onto a regular grid (refer to Allen et al. 2015 and Sharp et al. 2015 for details). The SAMI aperture of radius is approximately ∼7.5 arcsec with a sampling of $0''.5 \times 0''.5$ spaxels. The true spatial resolution is limited by the seeing, recorded as $\mathrm{FWHM}_{PSF} = 2.07$ arcsec (∼790 pc) for SAMI609396. SAMI observes in two spectral bands. The blue arm covers the observed wavelength range from 3750-5750 Å at low spectral resolution (R ∼ 1808, $\sigma v \sim 74$ km s^{-1}, at 4800 Å), while the red arm covers from 6300-7400 Å at medium resolution (R ∼ 4304, $\sigma v \sim 29$ km s^{-1}, at 6850 Å) (e.g., Zhou et al., 2017). For more detailed information on the SAMI survey and data products, the reader is referred to the above references.

Among nine SAMI galaxies in which we visually identified the presence of up to three auroral lines ([S II]$\lambda\lambda$4069, 76, [O III]λ4363 and [S III]λ6312), we highlight one notable case, SAMI609396 – a minor-merger system (Figure 4.1). The remainder of this paper is focused on this object. The list of SAMI galaxies we compiled with identifiable auroral line emission can be found in Appendix B.1.

4.2.1 SAMI609396

SAMI609396 (SDSS J114212.25+002004.0) is identified as a minor merger in the Sloan Digital Sky Survey (SDSS) images (Figure 4.1). The two merging galaxies are not deblended in the SDSS catalog with the merger system having a total r-band magnitude of 13.95. The SAMI input catalog gives the heliocentric redshift as $z = 0.01824$.

The merger signatures are evident from the colour difference and tidal tails. A visual inspection of the system shows one smaller galaxy exhibiting a strong blue colour, with a larger companion that is significantly redder (see Figure 4.1 middle panel). Spatially-resolved 1D spectra from the publicly available SAMI datacube show that the smaller

[2] https://sami-survey.org/abdr

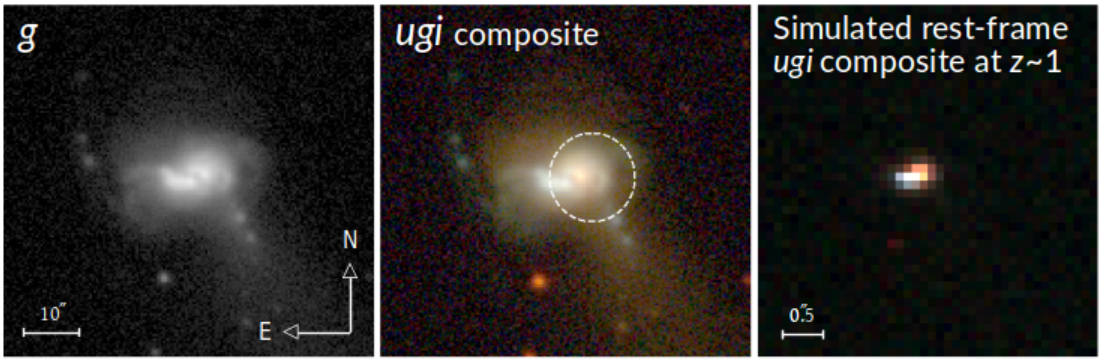

FIGURE 4.1: Left panel: g-band imaging of the SAMI609396 merger system from SDSS. Middle panel: *ugi* RGB composite of the system. Prominent auroral line emission is associated with SAMI609396B, the lower-left object exhibiting strong blue colour in the *ugi* composite. The white dashed circle in the middle panel shows the field of view observed by the SAMI IFU. The 10.″0 scale given for the g-band image applies also for the middle panel and corresponds to approximately 3.8 kpc in physical distance. Right panel: simulated rest-frame *ugi* colour composite after artificially redshifting the u-, g- and i-band imaging to $z \sim 1$. After redshifting, these bandpasses correspond approximately to *HST* filters ACS/F606W, ACS/814W, and WFC3/F160W. The pixel scale in the simulated image is 0.″1, similar to that of *HST*/WFC3. The simulated depth of the image is similar to observations in 3D-HST (Yuan et al., 2020).

galaxy in this system (SAMI609396B) is experiencing a burst of star formation associated with strong [O III]λ5007 emission lines (Equivalent Width (EW) $\sim$ 200 Å). Several prominent auroral emission lines ([S II]$\lambda\lambda$4069, 76, [O III]λ4363, and [S III]λ6312) are detected in SAMI609396B. Using spatially resolved star formation rate (SFR) maps and photometry from SAMI (Appendix B.2), we derive SFR and M_* estimates for SAMI609396B of 4.21 $\pm$ 0.30 $M_\odot$ yr^{-1} and $\log(M_*/M_\odot) = 9.18 \pm 0.05$. These values of SFR and M_* place SAMI609396B 1.3 dex above the local star formation 'main-sequence' (Renzini & Peng, 2015).

4.2.1.1 SAMI609396B properties in the context of high-redshift galaxies

A number of galaxy properties have been shown to evolve systematically with redshift including SFR (e.g. Speagle et al., 2014), metallicity (e.g. Zahid et al., 2013; Sanders et al., 2020a), ionisation parameter (Sanders et al., 2016), and nebular emission line ratios (e.g. Kewley et al., 2013; Steidel et al., 2014). Given that placing observational constraints on high-redshift galaxies is comparably much more challenging than for local galaxies, there has been interest in obtaining observational constraints for "high-redshift analogues" (e.g. Heckman et al., 2005; Cardamone et al., 2009; Green et al., 2014; Bian et al., 2016). These are galaxies at low-redshift with properties that emulate those

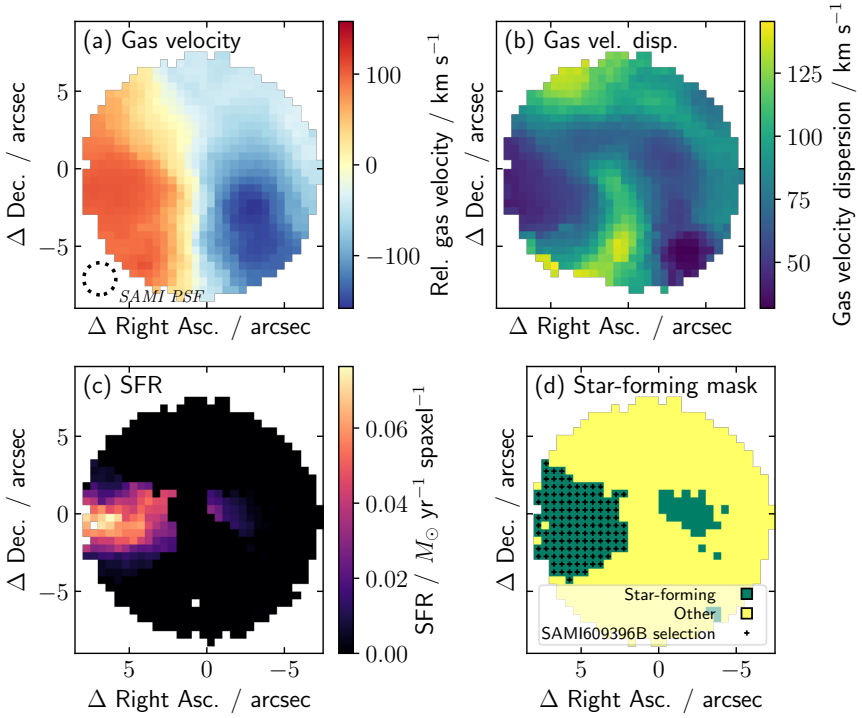

FIGURE 4.2: Publicly available value-added data products from SAMI DR2. Panel (a) Gas velocity from 1-component fitting. Panel (b) Gas velocity dispersion from 1-component fitting. Panel (c) Per spaxel star-formation rate (Medling et al., 2018). Panel (d) Star-forming mask. The large star-forming dominated region denoted with black '+' symbols in panel (d) is characterised by very high SFR, velocity dispersions of $\sim$30 – 80 km s^{-1}, and relative velocities of $\sim$100 km s^{-1} (in the scale of panel a). This region, designated SAMI609396B, is spatially associated with observed auroral lines and is the target of our investigation. The black dotted circle in panel (a) indicates the point-spread function measured for this SAMI observation and applies to all panels.

observed in high-redshift galaxies. Given the rarity of auroral emission lines in IFU data, we consider that SAMI609396B is worthy of a detailed study on its own. However, we also consider how its properties compare to those seen in high-redshift galaxies.

As outlined above in § 4.2.1, the SFR and M_* measurements for SAMI609396B are more than 1 dex above the local star-forming main sequence, more in line with values typical of galaxies at $z \gtrsim 1$. Global metallicity correlates positively with stellar mass at $z \sim 0$ (Mass-Metallicity Relation; refer to Maiolino & Mannucci 2019 and references therein), and at fixed stellar mass, metallicity is seen to decrease with increasing redshift (Zahid et al., 2013; Sanders et al., 2020a). According to a recent multi-diagnostic determination by Sanders et al. (2020a), galaxies of a mass comparable to SAMI609396B ($\log(M_*/M_\odot) \approx 9.18$) would have a median metallcity of $12+\log(O/H) = 8.55$ at $z \sim 0$,

$12+\log(O/H) = 8.26$ at $z \sim 2.3$, and $12+\log(O/H) = 8.17$ at $z \sim 3.3$. Absolute metallicity values for individual galaxies are notoriously difficult to determine and depend strongly on the calibration used (e.g. Kewley & Ellison, 2008). Although we do not take the step of applying the same metallciity calibration used by Sanders et al. (2020a), according to the metallicities we derive for SAMI609396B in § 4.4 we expect that the metallicity of SAMI609396B would likely fall somewhere between the median values expected from the $z \sim 0$ and $z \sim 2.3$ samples.

Ionisation parameters and electron densities in $z \sim 2.3$ galaxies have been shown to be systematically offset from local galaxies at fixed stellar mass (Sanders et al., 2016). Electron density is most commonly probed with the [S II] λ6716/λ6731 doublet ratio. MOSDEF galaxies at $z \sim 2.3$ were found by Sanders et al. (2016) to have a median [S II] doublet ratio of 1.13, corresponding densities of around 290 cm^{-3} in the S^+ zone of emitting nebulae, much higher than typical SDSS values ([S II] λ6716/λ6731 = 1.41; density of 26 cm^{-3}). We measure a global [S II] ratio of 1.29 for SAMI609396B, corresponding to a density of 118 cm^{-3}, placing SAMI609396B between the low- and high-redshift sample medians. Given the scatter about those median values in both the MOSDEF and SDSS samples (Figures 4 & 5 in Sanders et al. 2016), it is difficult to draw conclusions about how SAMI609396B compares to the two populations based on density. Using the O_{32}[3] strong-line ratio as a tracer for ionisation parameter, Sanders et al. (2016) found that, like SDSS galaxies, $z \sim 2.3$ MOSDEF galaxies show a trend of decreasing ionisation parameter with increasing stellar mass. They find the slope of this relation to be very similar to that of SDSS galaxies, however with a ∼0.6 dex offset toward higher O32 at fixed stellar mass in the $z \sim 2.3$ sample (Fig 8 in Sanders et al. 2016). Given the stellar mass derived for SAMI609396B ($\log(M_*/M_\odot) \approx 9.18$), SDSS galaxies have a median value of $O_{32} = -0.25$, while MOSDEF galaxies with comparable mass have much higher values ($O_{32} = 0.28$). We measure $O_{32} = 0.14$ for SAMI609396B, 0.39 dex higher than the median SDSS value and 0.14 dex below the median of $z \sim 2.3$ MOSDEF galaxies.

Recent studies have found an offset in the locus inhabited by high-redshift galxies on the $N2$-BPT diagram (Baldwin et al., 1981) with high-redshift galaxies exhibiting higher [O III]/Hβ at fixed [N II]/Hα (Kewley et al., 2013; Steidel et al., 2014). We find that BPT line ratios observed for SAMI609396B to be within the range of local galaxies.

[3] O_{32} = [O III] $\lambda\lambda$4959, 5007 / [O II] $\lambda\lambda$3726, 29 in this context

Our analysis of the spatially resolved BPT diagram of SAMI609396B is discussed in detail in § 4.5.

To summarise, we find that the physical properties (SFR and ISM conditions) of SAMI609396B tend to be offset from median $z \sim 0$ values, although are generally less extreme than $z \sim 2$ galaxies. In combination with the high EW([O III]), we consider that the physical properties of SAMI609396B might be analogous to intermediate-redshift ($0 < z \lesssim 1$) galaxies. Low-mass galaxies like SAMI609396B are extremely difficult to resolve at high redshift. To visually demonstrate what a system like SAMI609396 would look like at a higher redshift, we simulate the angular size and morphology of SAMI609396 at $z \sim 1$ using similar techniques to those detailed in Yuan et al. (2020). The redshifted morphology is presented on the right panel of Figure 4.1. In order to resolve a low-mass system like SAMI609396B at $z \sim 1$ with comparable physical resolution of SAMI, a minimal angular resolution of 0.1" is required. Such a fine resolution can be achieved either through ground-based adaptive optics or space instruments. The faintness of these low-mass systems also means the need for next-generation facilities such as JWST/NIRSpec and ground-based ELTs.

4.2.2 SAMI DR2: Value-added data products

SAMI DR2 includes a number of publicly available value-added data products, which we use to guide our initial understanding of the SAMI609396 system. Figure 4.2 shows publicly available maps for the gas velocity, gas velocity dispersion, and star-formation rate (Panels (a) – (c)) derived from 1-component fits.

Panel (d) of Fig 4.2 shows a star-formation mask, determined according to Kewley et al. 2006 based on BPT & VO87 diagnostic diagrams (Baldwin et al., 1981; Veilleux & Osterbrock, 1987), with green denoting spaxels passing selection as "star-formation dominated". Figure 4.2 shows that much of the SAMI field-of-view is dominated by emission from non-star-forming sources (yellow spaxels; "other"). The yellow spaxels have higher velocity dispersion compared with star-forming dominated regions, characteristic of emission from shock-heated gas. The BPT diagram and the origin of emissions in these regions are discussed further in §4.5.2.

TABLE 4.1: Global properties of SAMI609396B and its companion.

Right Ascension	$11^{h}\ 42^{m}\ 12.^{s}25$
Declination	$+00°\ 20^{m}\ 04.^{s}04$
z	0.01824
SAMI609396B:	
SFR ($M_\odot$ yr^{-1}) [a]	4.21 ± 0.30
$\log(M_*/M_\odot)$	9.18 ± 0.05
Companion:	
SFR ($M_\odot$ yr^{-1}) [a]	0.32 ± 0.08
$\log(M_*/M_\odot)$	9.88 ± 0.07

[a]SFR measurement for area within SAMI FoV (see Fig 4.1). This is best considered as a lower bound.

The prominent auroral line emission we identify is spatially associated with the large star-formation dominated region in the left-hand (eastern) portion of the star-formation mask. This region has a median rest-frame gas velocity of $v_{gas} \approx 100$ km s^{-1} (refer to scale in Fig 4.2), a velocity dispersion of range $\sigma_{gas} \approx 30 - 80$ km s^{-1}, and high a star-formation rate (median SFR surface density $\approx 0.97\ M_\odot$ yr^{-1} kpc^{-2}).

We designate this object as "SAMI609396B" and define its selection within the SAMI609396 datacube as including spaxels labelled as star-formation dominated with $v_{gas} > 0$, denoted by black '+' symbols in panel (d) of Fig 4.2. Global SFR and stellar mass estimates for SAMI609396B and its companion galaxy are provided in Table 4.1. Details of how these are derived are provided in Appendix B.2.

4.3 Spatially Resolved Electron Temperature

The electron temperature (T_e) and electron density (n_e) are fundamental physical parameters in understanding the emission line physics of ionized nebulae. Abundance measurements from collisionally excited lines in H II regions are very sensitive to these parameters. For this reason, chemical abundances derived following explicit measurements of T_e and n_e are generally used as a baseline calibration for understanding the chemistry of ionized nebulae (e.g. Maiolino & Mannucci, 2019).

This is generally achieved with the so-called "direct method" via measurement of an auroral emission line and a strong nebular line of the same ionic species. This is most

commonly applied to the O^{2+} ion using the [O III] λ4363/λ5007 ratio, which is primarily sensitive to T_e (its n_e dependence is minimal over the density range of typical H II regions). Within the typical rest-frame near-ultraviolet to near-infrared wavelength range observed for galaxies, auroral line ratios may be observable for a number of ionic species including O^+, N^+, S^{2+} and S^+, each of which probe different zones within the emitting H II regions according to the distribution of those ions within the nebular structure. Although we detect auroral lines from three ionic species in SAMI609396B ([S II], [O III] and [S III]), we are able to derive electron temperature for only the [O III] and [S II] ionisation zones as we lack the spectral coverage to measure the [S III] λ9069 and [S III] λ9531 strong lines required to derive T_e([S III]).

4.3.1 Auroral Emission Line Measurements

We derive flux maps for auroral lines from three ionic species ([S II] $\lambda\lambda$4069, 76, [O III] λ4363, and [S III] λ6312) identified in the SAMI609396 data cube, as the SAMI DR2 value-added data products do not include emission line maps for these fainter lines. We concomitantly re-derive strong emission line fluxes, rather than use SAMI DR2 emission line maps, ensuring self-consistency in our line ratio measurements. These flux maps are generated by applying standard methods to each spaxel, first fitting the stellar continuum, and then simultaneously fitting profiles to each emission line included in our analysis. Details of this spectral fitting are provided in Appendix B.3.

We obtain $S/N \sim 3-15$ in individual spaxels for each of [O III] λ4363, [S II] $\lambda\lambda$4069, 76 and [S III] λ6312 across the majority of the spatial region selected as SAMI609396B. We identify from visual inspection some degree of blending between [O III] λ4363 and a neighbouring faint [Fe II] emission line at λ4360, similar to that observed in other recent studies (e.g. Curti et al., 2017; Berg et al., 2020; Arellano-Córdova & Rodríguez, 2020). We find that the [O III] λ4363 line is brighter than the λ4360 feature by a factor of $\sim$2 and that with the spectral resolution of the blue arm of the SAMI spectrograph we are able to reliably recover the [O III] λ4363 flux. Our efforts to test the reliability of our [O III] λ4363 flux measurements are outlined in detail in Appendix B.3.3.

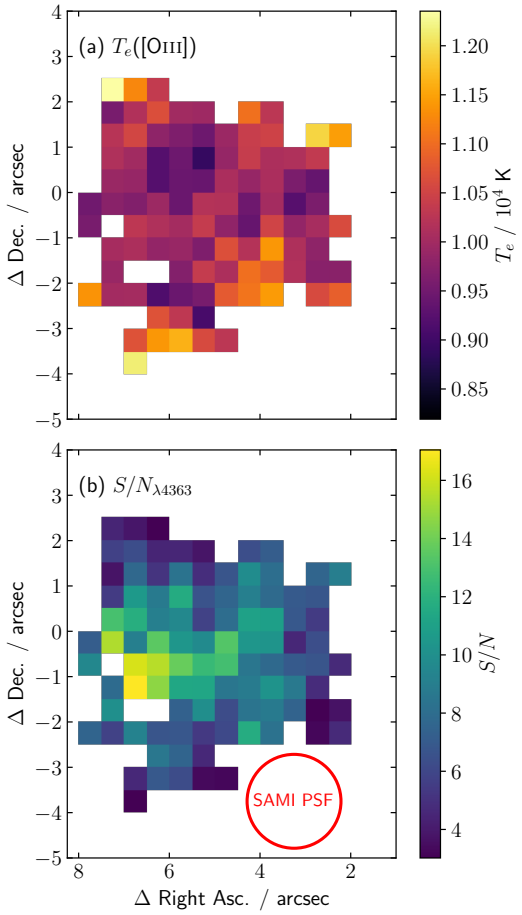

FIGURE 4.3: *Top*: Electron temperature map derived from the [O III]λ4363 / [O III]λ5007 ratio (see §4.3.2). *Bottom*: Measured signal-to-noise of [O III]λ4363 auroral line. The red circle depicts the FWHM PSF of this SAMI datacube and applies to both panels. The large scale spatial variations in T_e do not appear to correlate with [O III]λ4363 S/N.

4.3.2 [OIII] Electron Temperature

The emission line ratio most widely used to determine the electron temperature with the direct method is the [O III] λ4363 / [O III] λ5007 ratio. Despite the primary dependence of this [O III] ratio on temperature, the residual density dependence is often accounted for by measurement of a density sensitive line ratio, typically [S II] λ6716 / [S II] λ6731. Izotov et al. (2006) use relations derived for these aforementioned [O III] and [S II] line ratios (Equations (1) and (2) in that reference) in an iterative manner, solving simultaneously for T_e and n_e. This iterative approach is shared by the `getCrossTemDen` routine in the `PyNeb` package (Luridiana et al., 2015), which allows for a flexible array of temperature- and density-sensitive line ratios.

However, it is important to consider that neither temperature nor density is expected to be constant throughout H II regions. Additionally, emission from different ionic species may not be co-spatial. Certainly, [S II] emission is expected to arise from the outer regions of nebulae, thus densities measured from the [S II] line ratio do not necessarily provide a good indication of the density of the [O III] emission region (see Figure 2 in Kewley et al. 2019a).

Given these uncertainties, Nicholls et al. (2020) instead propose a simplified approach in which T_e is derived from an empirical relation of the auroral line ratio, derived from H II region modelling, forgoing any attempt to account for n_e, suggesting that any improvements in temperature insight are outweighed by uncertainties induced by density variations and lack of co-spatiality.

Given the ∼1 kpc spatial resolution of SAMI, we are unable to resolve individual H II regions, adding to the uncertainties described above. Thus, we use this simplified approach to derive our T_e from the [O III] λ4363 / [O III] λ5007 ratio according to the relation given in Nicholls et al. (2020). This relation is shown as Equation 4.1 here:

$$\log_{10}(T_e([\mathrm{OIII}])) = \frac{3.3027 + 9.1917x}{1.0 + 2.092x - 0.1503x^2 - 0.0093x^3} \tag{4.1}$$

where $x = \log_{10}(f_{4363}/f_{5007})$, with f_X referring to a line flux measurement of a collisionally excited line with rest-frame wavelength X Å, and T_e is in units of K. The derived [O III] temperature map for spaxels with [O III] λ4363 of S/N > 3 is shown in Figure 4.3.

4.3.3 [SII] Electron Temperature

In addition to T_e([O III]), spatially resolved measurements of the [S II] auroral lines allow us to measure T_e([S II]) from the [S II] $\lambda\lambda$4069, 76 / [S II] $\lambda\lambda$6716, 31 ratio.

Modelling indicates that at the low density limit ($1 < n_e < 50$ cm^{-3}), the residual density dependence of the [S II] $\lambda\lambda$4069, 76 / [S II] $\lambda\lambda$6716, 31 ratio is minimal. In contrast to the [O III] case, this [S II] temperature diagnostic is co-spatial with the [S II] density diagnostic, meaning that we are able to make a more reliable estimate of the density. The n_e values for SAMI609396 obtained with the [S II] λ6716 / [S II] λ6731 ratio (Eq. 3 in

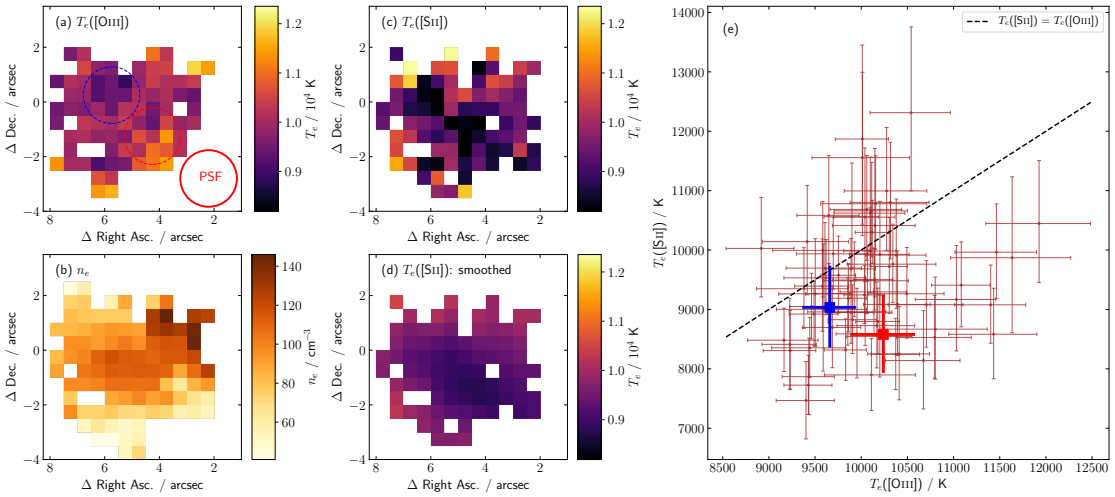

FIGURE 4.4: Comparison of $T_e([\mathrm{O\,III}])$ and $T_e([\mathrm{S\,II}])$ electron temperature values. *Panel (a)*: map of $T_e([\mathrm{O\,III}])$ values for spaxels with $S/N > 5$ for [O III] λ4363. Red circle labelled 'PSF' has diameter equal to the FWHM of the SAMI PSF for this observation and applies to panels (a–d). *Panel (b)*: electron density derived from the [S II] λ6716 / λ6731 ratio. *Panel (c)*: map of $T_e([\mathrm{S\,II}])$ values for spaxels with $S/N > 5$ for [S II] $\lambda\lambda$4069, 76. *Panel (d)*: map of $T_e([\mathrm{S\,II}])$ from panel (c) smoothed with a Gaussian filter. *Panel (e)*: Brown points show values of of $T_e([\mathrm{S\,II}])$ and $T_e([\mathrm{O\,III}])$ for individual spaxels with $S/N > 5$ on both auroral lines. Error bars shown reflect only measurement uncertainty and do not include associated modelling uncertainties. Temperatures derived for two mock apertures (indicated by blue and red dashed circles in panel a) are shown as the blue and red points in panel (e).

Proxauf et al., 2014b) are shown in Fig 4.4 Panel (b). We find the median electron density to be $\widetilde{n_e} = 92$ cm^{-3}. This value is above the [S II] low density limit, indicating that [S II] $\lambda\lambda$4069, 76 / [S II] $\lambda\lambda$6716, 31 will have a residual density dependence. Nonetheless, we derive the [S II] temperature with a similar approach to that outlined in §4.3.2 with a new rational polynomial fit to modelling data assuming a density of $n_e = 100$ cm^{-3}. This fit is given in Equation 4.2 where $x = \log_{10}[(f_{4069} + f_{4076})/(f_{6716} + f_{6731})]$ and T_e is in units of K.

$$\log_{10}(T_e([\mathrm{SII}])) = \frac{-0.08891 + 2.06354x + 3.38680x^2 + 0.10754x^3}{0.1 + 0.78000x + 0.94404x^2} \tag{4.2}$$

$T_e([\mathrm{S\,II}])$ values obtained for SAMI609396 are compared with $T_e([\mathrm{O\,III}])$ values in Figure 4.4. Panels (a) and (b) show maps of $T_e([\mathrm{S\,II}])$ and $T_e([\mathrm{O\,III}])$ respectively for spaxels where the relevant auroral line is detected with $S/N > 5$. Panel (e) shows the direct comparison of $T_e([\mathrm{S\,II}]))$ and $T_e([\mathrm{O\,III}])$ values on a spaxel-by-spaxel basis. We observe that a majority of points in panel (e) of Fig 4.4 lie below the line of $T_e([\mathrm{S\,II}]) = T_e([\mathrm{O\,III}])$

(i.e. higher $T_e([\mathrm{O\,III}])$ than $T_e([\mathrm{S\,II}])$). The large blue and red points in Fig 4.4 Panel (e) show derived $T_e([\mathrm{S\,II}])$ and $T_e([\mathrm{O\,III}])$ electron temperatures for two mock apertures which correspond to the regions shown as blue and red dashed circles in Panel (a). These aperture temperatures appear to indicate that $T_e([\mathrm{S\,II}])$ and $T_e([\mathrm{O\,III}])$ do not exhibit strong positive correlation across different spatial regions of SAMI609396B. The implications of this temperature relation for metallicity measurement are discussed further in §4.4.

4.4 Spatial Trends In Metallicity

In Section 4.3 we derived spatially resolved electron temperature (T_e) measurements. Here we use these T_e measurements to determine direct method oxygen abundances under three different sets of assumptions, showing that derived spatial variations in metallicity can be very sensitive to the assumed internal H II region temperature structure. Additionally we derive spatially resolved strong-line metallicities and discuss differences in observed spatial trends.

4.4.1 Direct Method Metallicity

Since the abundance of neutral oxygen (O^0) and oxygen in ionization states higher than O^{2+} is expected to be negligible in H II regions, we assume that the total oxygen abundance can be approximated as Equation 4.3:

$$\frac{O}{H} = \frac{O^+}{H^+} + \frac{O^{2+}}{H^+}. \tag{4.3}$$

We derive abundances of these two ionisation states of oxygen using the following analytic relations set out in Pérez-Montero (2017):

$$12 + \log\left(\frac{O^{2+}}{H^+}\right) = \log\left(\frac{f_{4959} + f_{5007}}{f_{H\beta}}\right) + 6.1868 + \frac{1.2491}{t(O^{2+})} - 0.5816 \cdot \log\Big(t(O^{2+})\Big) \tag{4.4}$$

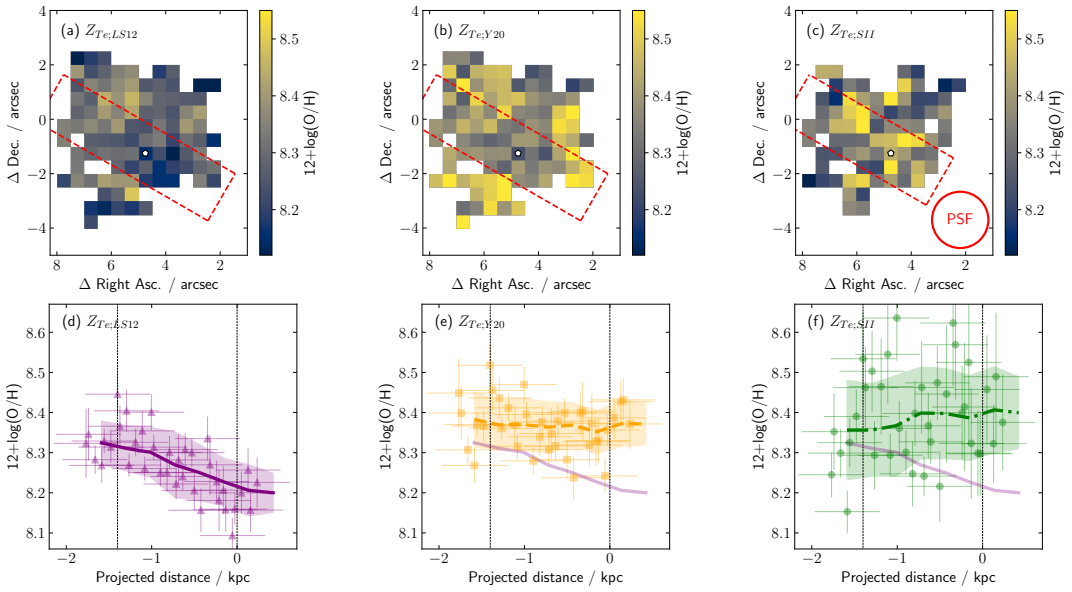

FIGURE 4.5: Observed spatial trends in direct method metallicity depend strongly on temperature structure assumptions. Direct method metallicity maps (panels a-c) and spatial metallicity trends (d-f) are shown for SAMI609396B under three different T_e([O II]) temperature assumptions. Panels (a, d) show $Z_{\mathrm{Te;\ LS12}}$: where T_e([O II]) is derived from T_e([O III]) via the relation of López-Sánchez et al. (2012) (Eq 4.6). Panels (b, e) show $Z_{\mathrm{Te;\ Y20}}$: derived as for $Z_{\mathrm{Te;\ LS12}}$ with the additional step of applying the empirical correction of Y20 based on $O32$. Panels (c, f) show $Z_{\mathrm{Te;\ SII}}$: metallicity is derived assuming T_e([O II]) = T_e([S II]). See §4.4.1 for details. Maps in panels (a-b) include spaxels with $S/N_{\lambda 4363} \geq 3$, while panel (c) additionally excludes spaxels with $S/N_{\lambda 4069} < 3$. The red circle in panel (c) shows the FWHM of the SAMI PSF and applies to panels (a-c). The dashed red rectangles in panels (a-c) span the region of highest S/N for the [O III]λ4363 line and defines the spatial region examined in panels (d-f). Panels (d-f) show individual points for which $S/N_{\lambda 4363} > 5$. Trend lines indicate running medians of the points shown. Vertical error bars on individual points reflect only measurement uncertainties and are dominated by auroral line measurements.

$$12 + \log\left(\frac{O^+}{H^+}\right) = \log\left(\frac{f_{3726} + f_{3729}}{f_{H\beta}}\right) + 5.887 + \frac{1.641}{t(O^+)} - 0.543 \cdot \log\Big(t(O^+)\Big) + 0.000114 \cdot n_e \quad (4.5)$$

where $t(O^{2+}) = T_e([\mathrm{O\,III}])/10^4$ K, $t(O^+) = T_e([\mathrm{O\,II}])/10^4$ K, n_e is the electron density measured by the [S II] λ6716 / λ6731 ratio, and f_X refers to a line flux measurement of the Hβ Balmer line or a collisionally excited line with rest-frame wavelength X Å. Deriving O^{2+}/H^+ in this way requires only [O III] $\lambda\lambda$4959, 5007 and Hβ emission line fluxes in addition to the T_e([O III]) values derived in § 4.3.2. On the other hand, the O^+/H^+ abundance from Eq. 4.5 calls for T_e([O II]), which we do not directly measure.

Additionally, O^+/H^+ has residual dependence on n_e, although we simply adopt the same fixed denisty $n_e = 100$ cm^{-3} used in the temperature calculations in §4.3.2. Note that our derived metallicity values vary by less than 0.01 dex with changes in adopted density, provided those are below $n_e < 200$ cm^{-3}.

Unlike T_e([O III]), we do not directly measure T_e([O II]), since we are unable to detect either the [O II] $\lambda\lambda$7319, 30 or [O II] $\lambda\lambda$2470+ doublets. A favourable alternative is to use temperatures derived from other ionic species, especially [N II] or [S III] , to probe the temperature structure (e.g. Berg et al., 2020). However, given the faintness of auroral lines it is common that an observation may enable measurement of only the [O III] temperature zone. In this scenario, a T_e([O II]) estimate can be obtained by adopting an empirical T_e([O II]) – T_e([O III]) relation, for which a number of calibrations exist (e.g. Izotov et al., 2006; López-Sánchez et al., 2012). Despite expanding the number of observations for which direct metallicities can be derived, Yates et al. (2020) (Y20 hereafter) find that using T_e([O II]) – T_e([O III]) relations can underestimate the direct metallicity by more than 0.5 dex for low-ionisation systems, highlighting the importance of constraining the internal temperature structure of H II regions where possible. Additionally, Y20 provide an empirical correction for this effect based on the [O III]/[O II] strong line ratio.

For this analysis, we determine our total oxygen abundance maps in three ways. Each differs in its approach to handling the O^+/H^+ abundance, while in all three cases the O^{2+}/H^+ abundance is determined from Eq 4.4 and our direct measurement of T_e([O III]). For the remainder of this paper, metallicities derived in these three ways will be abbreviated as $Z_{\mathrm{Te;\ LS12}}$, $Z_{\mathrm{Te;\ Y20}}$ and $Z_{\mathrm{Te;\ SII}}$ (where $Z = 12 + \log(O/H)$), described as follows:

1. $Z_{\mathrm{Te;\ LS12}}$: O^+/H^+ is determined using T_e([O II]) derived from T_e([O III]) using the relation outlined in López-Sánchez et al. (2012) (Eq 4.6).[4] This is the most commonly adopted method.

2. $Z_{\mathrm{Te;\ Y20}}$: As for $Z_{\mathrm{Te;\ LS12}}$, with the subsequent application of the Y20 empirical correction, based on [O III]/[O II] strong-line ratio (Eq 4.7). This is a relatively new correction and has not been widely implemented in literature yet.

[4]We note that alternative T_e([OII]) – T_e([OIII]) relations, including the equations from Izotov et al. (2006), do not significantly affect the metallicity morphology obtained for SAMI609396B.

3. $Z_{\mathrm{Te;\ SII}}$: O^+/H^+ is determined with $T_e([\mathrm{O\,II}])$ derived instead from $T_e([\mathrm{S\,II}])$ using the assumption $T_e([\mathrm{O\,II}]) = T_e([\mathrm{S\,II}])$. This is uniquely enabled by the detection of [S II] auroral lines in this study.

4.4.1.1 Empirical $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation

For $Z_{\mathrm{Te;\ LS12}}$ we adopt the $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation as calibrated by López-Sánchez et al. (2012), given in Equation 4.6:

$$T_e[\mathrm{O\,II}] = T_e[\mathrm{O\,III}] + 450 - 70 \cdot \exp\left[(T_e[\mathrm{O\,III}]/5000)^{1.22}\right] \tag{4.6}$$

Deriving $T_e([\mathrm{O\,II}])$ in this way and applying Equations 4.4–4.5 we obtain the total oxygen abundance map shown in panel (a) of Fig 4.5. The spatial structure of this map reflects that of the temperature map derived in Fig 4.3 and favours a strong trend in metallicity across the region of the highest signal-to-noise (Fig 4.5, panel d).

The measurement uncertainty is dominated by the flux uncertainty of the [O III] λ4363 emission line to the point where the measurement uncertainty contribution from the high S/N [O III], [O II] and Hβ strong lines can be ignored. We see no obvious correlation between the S/N of [O III] λ4363 and $T_e([\mathrm{O\,III}])$ (Fig 4.3). Increasing the minimum S/N cut on the [O III] λ4363 auroral line from $S/N > 3$ to $S/N > 8$ changes the median metallicity by less than 0.005 dex. Together, these give us confidence that observed spatial variations in metallicity are not artifacts from measurement noise, although the effects of modelling uncertainty are discussed over the coming sections.

4.4.1.2 Empirical O^+ Abundance Correction

Yates et al. (2020) provide an empirical correction based on the observed [O III]/[O II] line ratio given by Equation 4.7,

$$Z_{\mathrm{Te;\ Y20}} = Z_{\mathrm{Te;\ LS12}} - 0.71 \cdot (O32 - 0.29) \tag{4.7}$$

where $Z_{\mathrm{Te;\ Y20}}$ and $Z_{\mathrm{Te;\ LS12}}$ are corrected and uncorrected values of 12+log(O/H) respectively; $O32 = \log([\mathrm{O\,III}]\ \lambda\lambda 4959, 5007\ /\ [\mathrm{O\,II}]\ \lambda\lambda 3726, 9)$ and the correction is applied only when $O32 \leq 0.29$.

Values of $O32$ across SAMI609396B fall in the range for which this correction will be non-zero. Our direct metallicity map after Y20 correction is shown in Fig 4.5 Panel *(b)*. Spatial variations in the $O32$ ratio result in a flattening of the spatial trend after application of this correction.

We note that, in addition to the empirical correction described here ("Y20 correction"), Yates et al. (2020) also outlined a novel method for determining semi-direct metallicities ("Y20 method") in which $T_e([\mathrm{O\,II}])$ and metallicity are solved for simultaneously, rather than sequentially. This Y20 method then also requires subsequent application of the Y20 correction if $O32 \leq 0.29$, as above. Note that Figure 6 in Yates et al. (2020) shows that the abundance deficit at low O^{++}/O^{+}, which the Y20 correction adjusts for, is present to varying degrees for all $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relations considered in that work.

We find the Y20 method gives a two-valued solution for SAMI609396B which may require an additional prior to select the best metallicity solution. We found that applying the Y20 method as originally outlined favoured the lower value of these two solutions which yielded a gradient comparable to that obtained from our $Z_{\mathrm{Te;\ Y20}}$ approach here, albeit with a much lower normalisation (∼0.3 dex). We found that the normalisation of the upper-branch solution was in better agreement with our other determinations outlined here, however the spatial trend arising from this upper-branch solution is more difficult to interpret. Discussion of our implementation of the Yates et al. (2020) method and its two-valued nature is deferred to Appendix B.4.

4.4.1.3 $\mathbf{O^{+}}$ abundance with $T_e([\mathrm{S\,II}])$

The [S II] temperature samples a relatively narrow zone from the outer regions of nebulae and is consequently not widely used to constrain the temperature profile of emitting H II regions. However, Croxall et al. (2016) found general agreement of $T_e([\mathrm{S\,II}])$ with $T_e([\mathrm{O\,II}])$ and $T_e([\mathrm{N\,II}])$ in H II regions in NGC 5457. In the absence of the [S III] strong-lines, the [N II] auroral lines, or any other temperature probes, $T_e([\mathrm{S\,II}])$ affords our only direct probe of the internal temperature structure of H II regions in SAMI609396B.

We make the simplified assumption that $T_e([\mathrm{O\,II}]) = T_e([\mathrm{S\,II}])$ and update our total oxygen abundance using the measured $T_e([\mathrm{S\,II}])$ map (Fig 4.4 panel c) to re-derive our O^+/H^+ values. These updated oxygen abundances are shown in Fig 4.5 Panel (c), spanning a slightly smaller spatial extent due to the additional requirement of [S II] auroral line signal-to-noise. The spatial trend shown in Fig 4.5 Panel (f) is seen to be opposite of that in Panel (d) where O^+/H^+ was derived using an empirical temperature relation, albeit with a larger scatter.

This stark reversal can be explained by the $T_e([\mathrm{S\,II}]) - T_e([\mathrm{O\,III}])$ trends observed in Figure 4.4. Deriving $T_e([\mathrm{O\,II}])$ from a relation with $T_e([\mathrm{O\,III}])$ assumes that such a relation is fixed across the spatial region covered. This would mean that regions with elevated $T_e([\mathrm{O\,III}])$ would also show increased $T_e([\mathrm{O\,II}])$. However, the apertures plotted in panel (e) of Fig 4.4 (blue and red bold points) show that despite the increase in $T_e([\mathrm{O\,III}])$ from the 'blue' aperture to the 'red' aperture, measured $T_e([\mathrm{S\,II}])$ instead decreases (albeit with large uncertainties). This suggests the absence of a strong positive correlation between these temperatures across the spatial region and highlights the limitations of applying empirical temperature relations to measure spatial metallicity trends. This is likely driven by variations in the ionisation structure (i.e. O^{2+}/O^+ abundance ratio) and also explains the observed variations in $O32$ ratio that lead to the flattening of the spatial trend observed after applying the Y20 correction. We discuss this further in § 4.4.3.

4.4.2 Strong-Line Metallicity

In Figure 4.6 we compare four different strong-line metallicity maps with $Z_{\mathrm{Te;\,LS12}}$ and $Z_{\mathrm{Te;\,Y20}}$ direct method metallicity maps derived in §4.4.1. Strong-line metallicities are derived using a selection of widely strong-line diagnostics, defined in Equations 4.8 – 4.13:

$$N2O2 = \log_{10}\left([\mathrm{N\,II}]/[\mathrm{O\,II}]\right) \tag{4.8}$$

$$O32 = \log_{10}\left([\mathrm{O\,III}]/[\mathrm{O\,II}]\right) \tag{4.9}$$

$$R_{23} = \log_{10}\left(\frac{[\mathrm{O\,III}]\,\lambda4959 + [\mathrm{O\,III}]\,\lambda5007 + [\mathrm{O\,II}]}{H\beta}\right) \tag{4.10}$$

$$N2 = \log_{10}\left([\mathrm{N\,II}]/\mathrm{H}\alpha\right) \tag{4.11}$$

$$O3N2 = \log_{10}\left([\mathrm{O\,III}]/\mathrm{H}\beta\right) - N2 \tag{4.12}$$

$$N2S2H\alpha = \log_{10}\left([\mathrm{N\,II}]/[\mathrm{S\,II}]\right) - 0.264 \cdot N2 \tag{4.13}$$

where $[\mathrm{N\,II}] = [\mathrm{N\,II}]\,\lambda6583$, $[\mathrm{O\,II}] = ([\mathrm{O\,II}]\,\lambda3726 + [\mathrm{O\,II}]\,\lambda3729)$, $[\mathrm{S\,II}] = ([\mathrm{S\,II}]\,\lambda6716 + [\mathrm{S\,II}]\,\lambda6731)$, and $[\mathrm{O\,III}] = [\mathrm{O\,III}]\,\lambda5007$ unless otherwise specified. We use strong-line calibrations based on a mixture of theoretical and observational calibrations, outlined as follows:

- **N2O2**: We use the theoretical calibration provided in Kewley et al. (2019a) to solve iteratively for metallicity and ionisation parameter using the $N2O2$ (Eq. 4.8) and $O32$ (Eq. 4.9) diagnostic line ratios.
- **$\mathbf{R_{23}}$**: We use the calibration provided by Curti et al. (2020a) based on direct method measurements of stacked SDSS galaxies. The R_{23} ratio (Eq. 4.10) is two-valued with a turnover at around $12+\log(O/H) = 8.1$. Using $N2$ (Eq. 4.11) to distinguish between high- and low-metallicity branches, we find $N2 > -1.0$ across the extent of SAMI609396B, prompting us to consider only the high-metallicity branch.
- **O3N2**: Calibration based on large compilation of T_e measurements in H II regions from Marino et al. (2013).
- **N2S2Hα**: This diagnostic was proposed by Dopita et al. (2016) based on predictions from photoionisation modelling. We adopt the calibration presented therein.

The colour maps shown in Fig 4.6 are shown with different normalisation so as to visualise any spatial trends in metallicity in each diagnostic, setting aside the expected discrepancies in normalisation between alternative diagnostics (e.g. Kewley & Ellison, 2008).

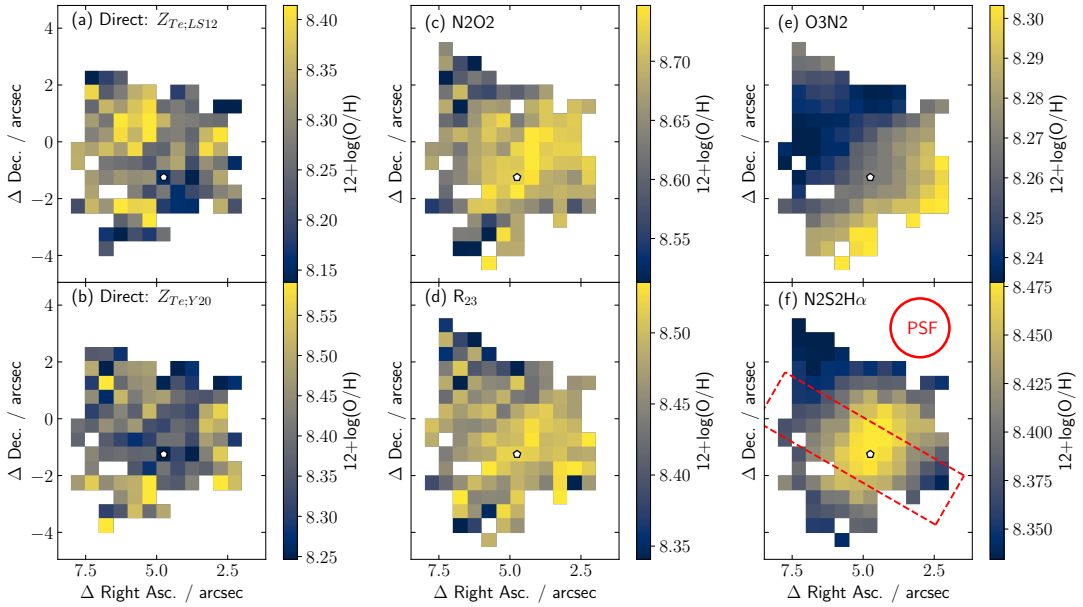

FIGURE 4.6: Direct method and strong-line oxygen abundance maps for the star-formation selected region corresponding to SAMI609396B. *Panel (a)*: Direct method metallicity using T_e values derived from [O III] λ4363 / λ5007 ratio (see §4.4.1). *Panel (b)*: Direct method metallicity after applying the empirical correction of Yates et al. (2020) (see §4.4.1.2). *Panel (c)*: iterative solution for metallicity, solved simultaneously for metallicity with $N2O2$ and ionisation parameter with $O32$ using calibrations from Kewley et al. (2019a). *Panel (d)*: metallicity from R_{23} strong-line diagnostic using calibration from Curti et al. (2020a). *Panel (e)*: metallicity derived from $O3N2$ using calibration from Marino et al. (2013). *Panel (f)*: metallicity derived from the $N2S2H\alpha$ diagnostic as outlined in Dopita et al. (2016). The peak i-band flux from SDSS imaging is marked in each panel with a white pentagon. FWHM of the spatial PSF is shown by the red circle in panel (f). The slit shown in panel (f) spans the region of highest S/N for the [O III] λ4363 line and is examined in detail in Fig 4.7

Indeed, even after applying the Y20 correction, the median direct method metallicity ($\widetilde{Z}_{\mathrm{Te;\ Y20}} = 8.40$) is still nearly 0.3 dex lower than that of the theoretically calibrated $N2O2$ diagnostic ($\widetilde{Z}_{\mathrm{N2O2}} = 8.68$). This difference is consistent with previous work which has shown systematic offset between metallicities derived from $N2O2$ using theoretical and empirical calibrations (Bresolin et al., 2009; Bresolin & Kennicutt, 2015).

4.4.3 Is The Metallicity Gradient Positive Or Negative?

While it is widely known that different metallicity measurement techniques often disagree in normalisation, one would hope that at a minimum two methods should agree on the ranked order of metallicities they measure. It is immediately striking from Fig 4.6 that even qualitative spatial trends in metallicity are very sensitive to the adopted diagnostic.

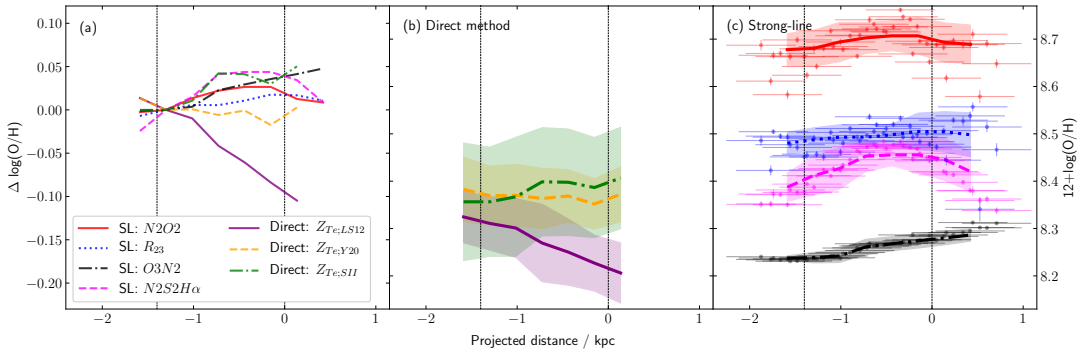

FIGURE 4.7: Spatial trend in metallicity along a mock slit for seven different strong-line and direct method metallicity measurement techniques. Panel (c) shows individual spaxels and running median trends measured with $N2O2$ (red), $O3N2$ (black), R_{23} (blue) and $N2S2H\alpha$ (magenta). Colour coding is as indicated in the legend in panel (a). More details on these strong-line metallicities can be found in § 4.4.2. Panel (b) reproduces trend lines for three different direct method assumptions from Fig 4.5 for ease of comparison. Panel (a) renormalises each of these seven trend lines to show metallicity deviation. The horizontal axis is zeroed at the adopted core of SAMI609396B, taken as the location of the peak in i-band flux from SDSS imaging. Vertical error bars show the measurement uncertainty carrying through from emission line measurements. Horizontal error bars indicate the FWHM of the spatial PSF of the SAMI observation in terms of physical distance.

Figure 4.7 illustrates these spatial trends as a 1D projection. Given the disturbed morphology of SAMI609396B, we do not formally define a metallicity gradient, but instead examine 1D spatial trends along the mock slit shown in Fig 4.5 panels (a-c) and Fig 4.6 panel (f). This slit encompasses the region of highest emission line signal-to-noise and approximately corresponds to the region of highest g-band flux (Fig 4.1).

Panel (a) of Fig 4.7 shows the running medians in metallicity with projected distance along this mock slit for all four strong-line methods described in § 4.4.2 as well as the three different direct method assumptions outlined in § 4.4.1. The distance axis has been zeroed at the location of peak i-band flux from SDSS imaging which we adopt as the core of SAMI609396B. Each trend line has been renormalised relative to the metallicity at $r = -1.3$ kpc. We renormalise at this projected distance rather than the core as the three direct method approaches show best agreement in this spatial region (Fig 4.7 Panel b). In particular, the Y20 empirical corrections are smallest in this region.

Most striking in Fig 4.7 Panel (a) is the clear discrepancy between the $Z_{\mathrm{Te;\,LS12}}$ direct method and all other methods. The $Z_{\mathrm{Te;\,LS12}}$ method favours a strong trend of decreasing metallicity left-to-right from negative projected distance toward the core. Strong-line methods show an opposite trend, with metallicity increasing in the same direction albeit

with less overall deviation from uniform. As outlined in § 4.4.1, we find that the $Z_{\mathrm{Te;\ Y20}}$ and $Z_{\mathrm{Te;\ SII}}$ direct methods both show a much flatter metallicity trend than the $Z_{\mathrm{Te;\ LS12}}$ method, and are in better agreement with strong-line methods.

Given that strong-line methods have their own unsettled systematic uncertainties (Section 4.5.1), we do not assess the absolute correctness of 'gradients' derived from each method. Instead, we discuss below the physical reason for why the gradient from the $Z_{\mathrm{Te;\ LS12}}$ method is at odds with $Z_{\mathrm{Te;\ Y20}}$ and $Z_{\mathrm{Te;\ SII}}$ and the strong line methods.

4.4.4 O^{2+}/O^{+} abundance ratio variation

We attribute the cause of the discrepancy between $Z_{\mathrm{Te;\ LS12}}$ and other methods to variations in the O^{2+}/O^{+} abundance ratio, causing deviations from the fixed $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation adopted by $Z_{\mathrm{Te;\ LS12}}$. Figure 4.8 shows separate O^{+}/H^{+} and O^{2+}/H^{+} abundance maps, derived using $T_e([\mathrm{S\,II}])$ and $T_e([\mathrm{O\,III}])$ respectively, with panel (a) showing elevated O^{+}/H^{+} in the core region (lower-right; corresponding to Projected Distance ≈ 0 kpc in horizontal scale of Fig 4.7).

A bulk change in the ionisation structure of H II regions across SAMI609396B such as this would cause measured temperatures to deviate from the $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation from López-Sánchez et al. (2012) (Eq 4.6).[5] In § 4.3.3 we noted that $T_e([\mathrm{S\,II}])$ and $T_e([\mathrm{O\,III}])$ derived for two mock apertures indicated the absence of a strong positive correlation between $T_e([\mathrm{S\,II}])$ and $T_e([\mathrm{O\,III}])$ (Fig 4.4 Panel e). In particular, lower $T_e([\mathrm{S\,II}])$ values obtained in the core region leads to systematically higher O^{+} abundance measurements in $Z_{\mathrm{Te;\ SII}}$ than $Z_{\mathrm{Te;\ LS12}}$, driving the apparent reversal in the measured total oxygen abundance gradient.

Recently, Yates et al. (2020) observed that for $\log(O^{2+}/O^{+}) \lesssim 0.0$, "semi-direct" metallicities (that is, metallicities in which $T_e([\mathrm{O\,III}])$ has been directly measured, but $T_e([\mathrm{O\,II}])$ has been indirectly determined using an assumed $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation) underestimated the total metallicity by up to ∼0.5 dex compared with metallicities derived using direct measurements of both $T_e([\mathrm{O\,II}])$ and $T_e([\mathrm{O\,III}])$. This effect also correlates with the [O III]/[O II] strong-line ratio, motivating the Y20 correction for observations with $\log([\mathrm{O\,III}]\,\lambda\lambda 4959, 5007\ /\ [\mathrm{O\,II}]\,\lambda\lambda 3726, 9) \leq 0.29$.

[5] Or, indeed, any fixed monotonic relation assumed between $T_e([\mathrm{O\,II}])$ and $T_e([\mathrm{O\,III}])$.

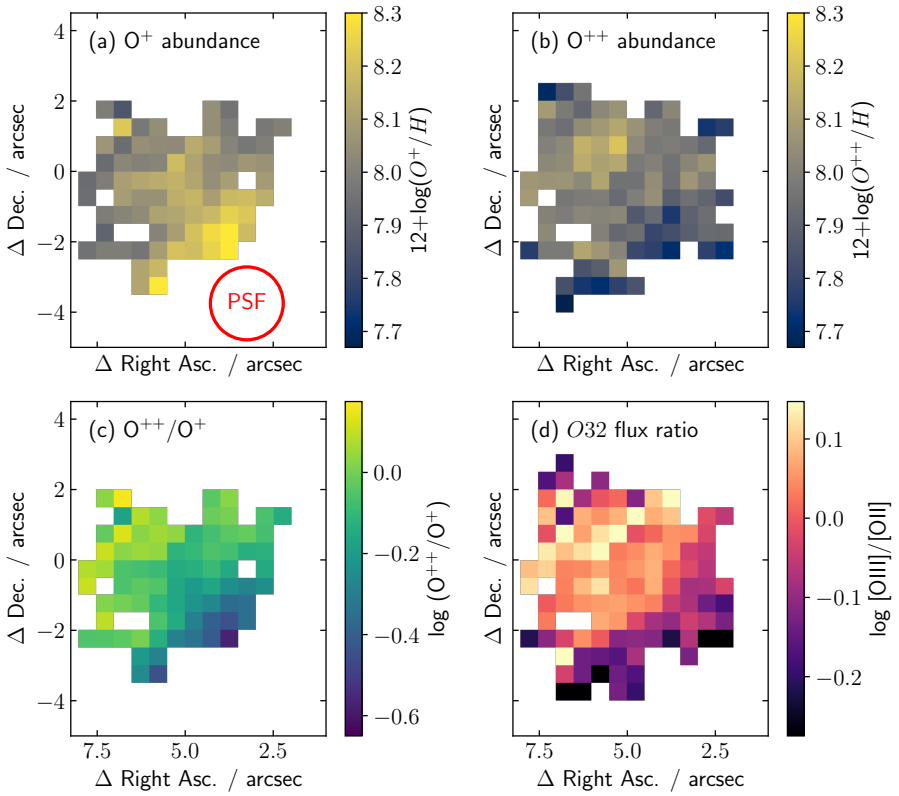

FIGURE 4.8: Map of derived O^{++}/O^{+} abundance ratio for SAMI609396B. Panel (a): O^{+}/H^{+} abundance derived from Eq 4.5 using T_e[OII]=T_e[SII]. Panel (b): O^{++}/H^{+} abundance derived from Eq 4.4 with direct T_e[OIII] measurement. Note that, unlike in Figures 4.5 & 4.7, abundance maps derived here use maps of T_e[SII] and T_e[OIII] that have been smoothed by a gaussian filter (FWHM set to measured PSF) to aid in the visual representation of spatial trends. Panel (c): O^{++}/O^{+} abundance ratio. O^{+} provides a larger contribution to the total oxygen abundance across the majority of SAMI609396B ($\log(O^{++}/O^{+}) < 0$). Direct metallicities evaluated adopting an assumed T_e[OII] – T_e[OIII] relation (e.g. $Z_{\mathrm{Te;\,LS12}}$ in this paper) can underestimate the total oxygen abundance by up to ∼0.5 dex in this low ionisation regime (see Fig 7 in Y20). Panel (d): Observed $O32$ strong-line ratios (Eq 4.9) appear to correlate with the O^{++}/O^{+} abundance ratio when O^{+}/H^{+} abundance is derived in this way. Regions of lowest $O32$ correspond to the highest level of correction according to Y20 correction (see § 4.4.1.2).

Figure 4.8 shows that O^{2+}/O^{+} abundance ratios in SAMI609396B largely fall below $\log(O^{2+}/O^{+}) \lesssim 0.0$, inside the range highlighted in Y20 as giving rise to deficits in the total oxygen abundance when "semi-direct" methods are used. Furthermore, a spatial trend in O^{2+}/O^{+} abundance ratio can be seen in panel (c) of Figure 4.8, with lower O^{2+}/O^{+} in the lower-right regions of SAMI609396B. Y20 found that the "semi-direct" abundance deficit is more pronounced at lower values of O^{2+}/O^{+}. From this, we reason that it is likely that $Z_{\mathrm{Te;\,LS12}}$ underestimates the total oxygen abundance across the majority of SAMI609396B. In particular, the lower O^{2+}/O^{+} seen in the core

TABLE 4.2: Input parameters for basis points shown in Figure 4.9.

		$Z/Z_\odot$	$\log(q)$	κ
H II region [a]		1.0	7.75	50
	v (km s^{-1})	$Z/Z_\odot$	n (cm^{-3})	B (μG)
Fast shock [b,d]	250	1.0	10	10
Slow shock [c,d]	160	1.0	1000	6.1

[a]Dopita et al. (2013); [b]Allen et al. (2008)
[c]Sutherland & Dopita (2017)
[d]Shock basis points include 50% contribution from pre-cursor

of SAMI609396B indicate that the systematically lower metallicities obtained in the core versus higher radius for $Z_{\mathrm{Te;\,LS12}}$ (panel d of Figure 4.5) can be explained by this semi-direct abundance deficit being amplified in the core region.

By not appropriately accounting for this trend, when applying the $Z_{\mathrm{Te;\,LS12}}$ method the O^{2+}/O^{+} abundance ratio trend instead masquerades as the trend in total oxygen abundance seen in Fig 4.5 & 4.7.

4.5 Discussion

4.5.1 Finer metallicity trends from strong lines

The measurement uncertainties on direct method metallicities for SAMI609396B are too large to be used for anything more than the bulk trend. While the strong-line methods show general agreement when considered in this bulk fashion, deviations exist in the finer details of their spatial trends (Fig 4.6 panels c-f & Fig 4.7 panel c). Most notable is the tendency of $O3N2$ to continue to increase beyond the core ($r > 0$ kpc in Fig 4.7), out to the boundary of the star-formation selected region. While other strong line methods, especially $N2S2H\alpha$ and $N2O2$, favour a peak in metallicity around $r = -0.6$ kpc and decreasing past the core and beyond. We explore the possibility of this tension as arising from contaminating emission from non-star-forming sources below in § 4.5.2.

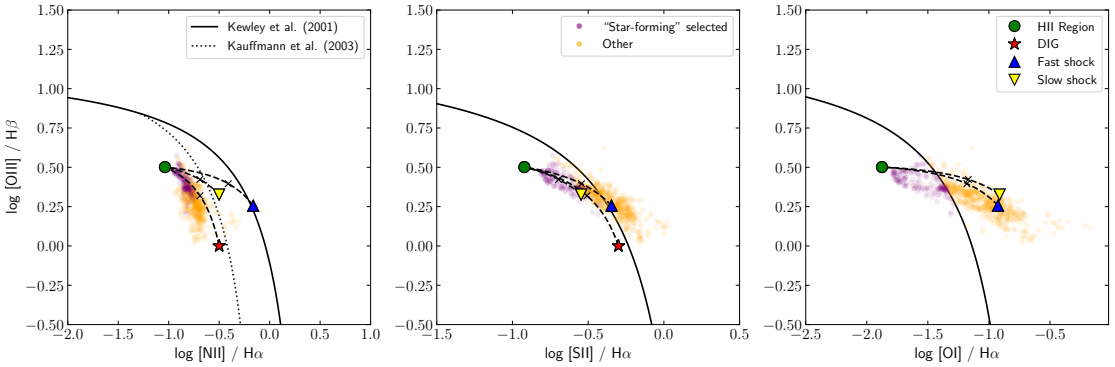

FIGURE 4.9: BPT & VO87 diagnostics diagrams for SAMI609396B. Line ratios for individual spaxels are shown as orange and purple points. Kewley et al. (2001) and Kauffmann et al. (2003) demarcation lines are shown as solid and dotted grey lines respectively. Purple points denote spaxels below these demarcation lines in each panel. Solid shapes are basis points predicted from photoionisation modelling for H II regions (Dopita et al. 2013; green circles), fast shocks (Allen et al. 2008; blue triangle) and slow shocks (Sutherland & Dopita 2017, Dopita & Sutherland 2017; yellow inverted triangle) according to the model parameters given in Table 4.2. The DIG basis point (red stars) is adopted as the peak region of strong-line ratios from the 10% lowest surface brightness spaxels in the Zhang et al. (2017) MaNGA sample (Sanders et al., 2017). Black dashed lines indicate fractional mixing sequences between these basis points.

4.5.2 Dissecting the Emission Line Excitation Mechanisms on the BPT Diagram

Gas-phase metallicity studies such as this aim to determine abundances of nebulae photoionised by recently formed O- and B-type stars (H II regions). However, emission from other sources including active galactic nuclei (AGN), shock-heated gas (shocks), and diffuse ionised gas (DIG), may contribute significantly to an observed extragalactic emission spectrum. Since each of these sources exhibit characteristically different emission spectra, inference of the properties of ionised gas from an emission line spectrum requires knowledge (or an assumption) of the excitation mechanism causing the emission.

Different excitation sources are generally distinguished with BPT or VO87 diagnostic diagrams which compare [O III]/Hβ to each of [N II]/Hα, [S II]/Hα, and [O I]/Hα (Baldwin et al., 1981; Veilleux & Osterbrock, 1987). Demarcation lines that separate H II regions from other sources of emission have been derived from photoionisation modelling (Kewley et al., 2001) and from large samples of observational data (Kauffmann et al., 2003). These can be used to exclude observations which are dominated by emission sources other than H II regions.

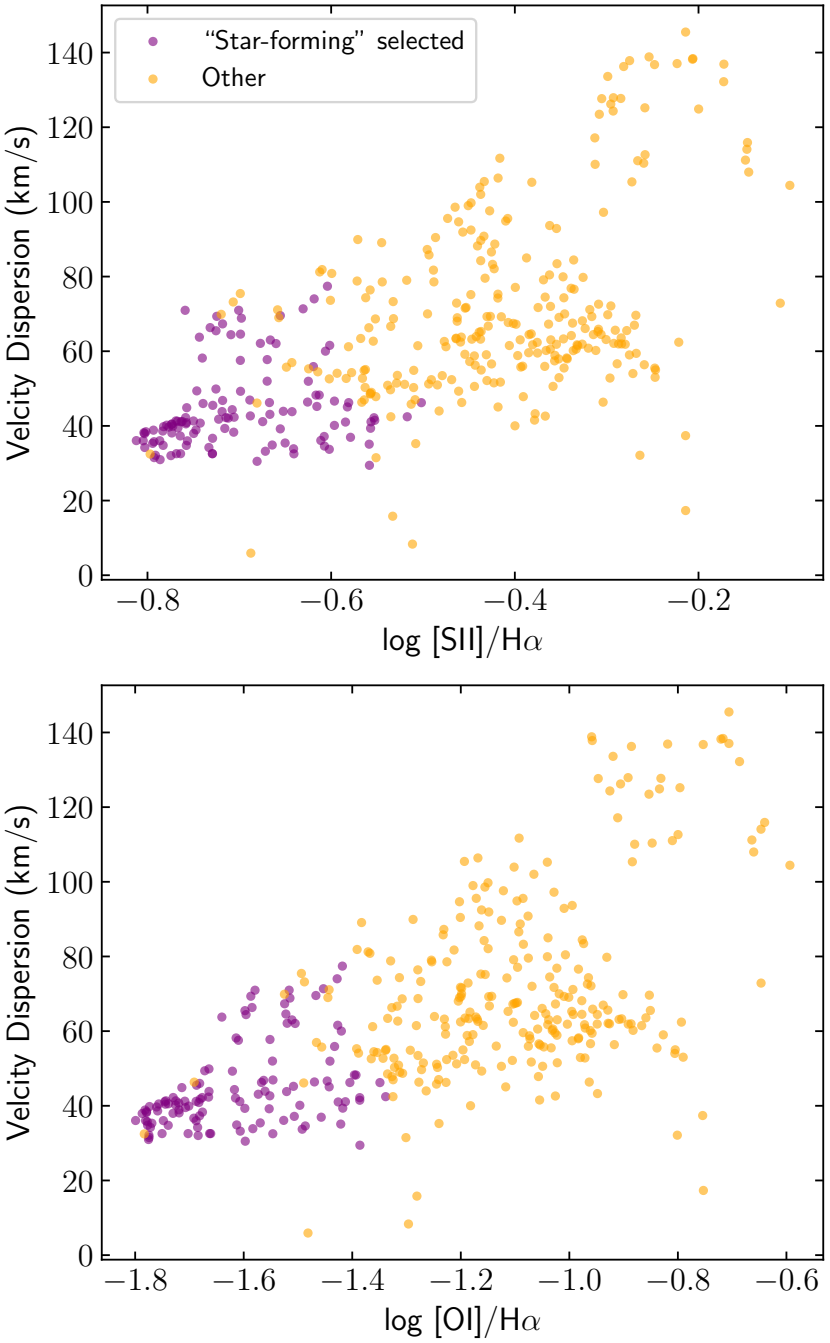

FIGURE 4.10: [S II]/Hα and [O I]/Hα diagnostic line ratios plotted against velocity dispersion for the full SAMI609396 field of view. The positive correlation observed between each of these diagnostic line ratios and velocity dispersion indicates the presence of shocks. The line ratio shown on the horizontal axis is [S II] $\lambda\lambda$6716, 31 / Hα for the top panel and [O I] λ6300 / Hα in the bottom panel. Colour coding is as for Figure 4.9. Emission line fluxes and velocity dispersions shown in this figure are from 1-component fits provided in the SAMI DR2 value-added data products.

Of course, the presence of one emission source in an observation does not preclude the presence of any others. Indeed a so-called "mixing sequence" is often observed on diagnostic diagrams, spanning the regions between the loci inhabited by H II regions and those of other ionizing sources. Global spectra residing along this sequence are best explained as galaxies for which the global spectrum contains emission from both H II regions and either AGN or shocks, with the position along this mixing sequence determined by the relative proportion of each of these sources of emission. Further, when observations are made with IFU spectroscopy, mixing sequences can be spatially resolved within individual galaxies (Ho et al., 2014; Davies et al., 2014a,b, 2016, 2017; Jones et al., 2017; Zhang et al., 2017; D'Agostino et al., 2018) due to differing spatial distributions of emission sources within these galaxies.

Figure 4.9 shows diagnostic line ratios for individual spaxels from SAMI DR2 single component emission line fits over the full extent of the SAMI609396 merger system. Purple points are those which pass the H II region Kewley et al. (2001) selection criteria in all three panels. The spatial region selected as SAMI609396B analysed in this paper is a subset of these purple points (refer to Fig 4.2 for SAMI609396B spatial selection).

Overplotted on Fig 4.9 are basis points predicted from photoionisation modelling for H II regions (Dopita et al. 2013; green circles), fast shocks (Allen et al. 2008; blue triangle) and slow shocks (Sutherland & Dopita 2017, Dopita & Sutherland 2017; yellow inverted triangle) as well as observed loci of DIG-dominated regions (Zhang et al. 2017; Sanders et al. 2017; red star). The adopted model parameters for each of these basis points are summarised in Table 4.2. Note that the shock model basis points include a contribution from precursor emission. We assume a 50:50 contribution from the shock and precursor. H II region model parameters are based on metallicity and ionisation parameter values obtained from $N2O2$ and $O32$ line ratios (see § 4.4.2). Shock model parameters are difficult to constrain as they are degenerate with fractional contribution and spatial variations, not to mention the large modelling uncertainties. Selected shock velocities (Table 4.2) broadly reflect velocity dispersions observed in SAMI609396 (see Fig 4.2 & 4.10) and were chosen on the basis of how well they visually reproduced the individual points in Figure 4.9. Black dashed lines show mixing models between H II regions and each of these other emission sources. These lines indicate the sequence that arises by varying in the fractional contribution between the two fixed basis points. The mid-point of each sequence is labelled with a black cross.

In addition to emission line ratios, velocity dispersion is a useful tool for identifying the presence of shocks. Emission from shocks often shows a positive correlation between velocity dispersion and [S II]/Hα or [O I]/Hα diagnostic line ratios (Ho et al., 2014), while DIG emission will not yield such a correlation. In Figure 4.10, [S II]/Hα and [O I]/Hα emission line ratios from SAMI609396 are plotted against measured velocity dispersion, supplementing our BPT and VO87 diagrams. Figure 4.10 shows that both [S II]/Hα and [O I]/Hα ratios are positively correlated with velocity dispersion in SAMI609396. While emission line ratios alone cannot definitively distinguish between emission from shocks and DIG (Fig 4.9), on the basis of Figure 4.10 we conclude that the dominant source of non-star-forming emission observed in the SAMI609396 data cube is shock-heated gas.

4.5.2.1 Effect of contaminating emission

Given the limited ($\sim$kpc) spatial resolution of SAMI, some amount of contamination from non-star-forming emission sources is inevitable, despite limiting our analysis to the region of nominally star-forming dominated emission. Sanders et al. (2017) showed that contamination from DIG can lead to discrepancies in measured metallicity of up to $\sim$0.3 dex. In resolved studies, Poetrodjojo et al. (2019) found that the inclusion of DIG in metallicity gradient measurements affects all diagnostics to varying degrees.

Of particular concern to establish the robustness of gradient studies is the presence of significant systematic variation in the relative contribution of H II region and non-star-forming emission. This has the potential to affect the inference on spatial metallicity trends. Figure 4.9 suggests that spaxels in this star-forming selected region may form the beginning of a spatial mixing sequence, perhaps indicating existence of spatial variations in the fractional contribution of shock emission to the total emission. Given the multiple ways metallicities from different diagnostics can be affected by contaminating emission, these variations could help to explain differences in the apparent metallicity trends recovered.

Line ratios plotted in Figures 4.9 & 4.10 support our assumption that the "star-forming" selected spaxels associated with SAMI609396B are indeed dominated by emission from H II regions. However, it should be considered that even in regions with emission "dominated" by H II regions, some amount of non-star-forming emission will invariably be

present. In particular, the mixing sequences shown as black dashed lines in Figure 4.9 highlight that there is room for variation in the relative contribution of different emission sources without moving outside the scope of what can be considered "dominated" by H II regions. A quantitative assessment of this effect is beyond the scope of this paper, but we note that variable contributions of non-star-forming emission in IFU observations of galaxies has the potential to affect measured trends in gas-phase abundances.

In § 4.4 we showed that, aside from the $Z_{\mathrm{Te;\,LS12}}$ application of the direct method, our metallicity measurements favour a flattened metallicity gradient. This flat gradient is likely due to the effects of the merger, which are known to produce flattened metallicity gradients due to strong inflows of pristine galaxies from the outskirts of galaxies (e.g. Kewley et al., 2010). The measured gradient may be affected by the presence of shocks, however given that these metallicities were derived using a relatively small subset of the mixing sequence seen in Fig 4.9 (i.e. the purple points) the effect of this contribution is likely not too significant.

4.6 Conclusion

Following a search of the SAMI Galaxy Survey Data Release 2 Public Data, we identified SAMI609396B, an interacting galaxy showing high S/N, spatially-resolved detections of three auroral lines: [O III] λ4363, [S II] $\lambda\lambda$4069, 76 and [S III] λ6312. The source also has properties that make it a good candidate for a local analog of high redshift galaxies, in particular for its combination of moderate stellar mass, disturbed morphology and elevated specific star formation rate (see § 4.2.2 & Appendix B.2).

We use [O III] and [S II] auroral-to-strong line ratios to derive spatially resolved electron temperature measurements for two sub-regions within the emitting H II regions (T_e([O III]) and T_e([S II])). Our results indicate the absence of a strong positive correlation between the T_e([S II]) and T_e([O III]) temperatures across different spatial regions in SAMI609396B. Instead, Figure 4.4 shows T_e([S II]) and T_e([O III]) appearing to trend in opposite directions between two apertures. This deviates from the common assumption of a fixed positive monotonic relation between these different temperatures.

Our $T_e([\mathrm{O\,III}])$ measurements allow for direct method O^{2+}/H^+ abundance measurements. We then derive direct method total oxygen abundances under three different treatments of the O^+/H^+ abundance:

1. $Z_{\mathrm{Te;\,LS12}}$: $T_e([\mathrm{O\,II}])$ is assumed from $T_e([\mathrm{O\,II}]) - T_e([\mathrm{O\,III}])$ relation (López-Sánchez et al., 2012).

2. $Z_{\mathrm{Te;\,Y20}}$: As for $Z_{\mathrm{Te;\,LS12}}$, with additional Y20 empirical correction, based on [O III]/[O II]strong-line ratio.

3. $Z_{\mathrm{Te;\,SII}}$: $T_e([\mathrm{O\,II}])$ adopted as $T_e([\mathrm{O\,II}]) = T_e([\mathrm{S\,II}])$.

We show that the disagreement between spatial metallicity trends returned by these methods is pronounced. $Z_{\mathrm{Te;\,LS12}}$ favours a strong spatial trend with lower total oxygen abundances being measured in the core, while $Z_{\mathrm{Te;\,Y20}}$ and $Z_{\mathrm{Te;\,SII}}$ instead suggest a flatter spatial trend, if anything perhaps opposite to the $Z_{\mathrm{Te;\,LS12}}$ trend. We conclude that the cause of this disagreement is variation in the O^{2+}/O^+ abundance ratio causing deviations from the assumed $T_e([\mathrm{O\,II}]) - T_e([\mathrm{O\,III}])$ relation. Accordingly, $Z_{\mathrm{Te;\,LS12}}$ results in systematically lower O^+ abundances across the whole of SAMI609396B than those of $Z_{\mathrm{Te;\,SII}}$. This gives rise to an apparent metallicity gradient as the effect is not spatially uniform: O^+ abundance is particularly elevated in the core when probed by $Z_{\mathrm{Te;\,SII}}$. The measured variation in the O^{2+}/O^+ abundance ratio correlates with variations in the [O III]/[O II] strong line ratio. Thus, applying the empirical correction from Yates et al. (2020) ($Z_{\mathrm{Te;\,Y20}}$) results in a trend more in line with $Z_{\mathrm{Te;\,SII}}$. Additionally, we derive metallicity with four strong-line diagnostics (R_{23}, $N2O2$, $O3N2$ and $N2S2H\alpha$) using a mixture of observation- and theory-based calibrations. Spatial trends recovered by these strong-line methods again favour opposite trends to that of $Z_{\mathrm{Te;\,LS12}}$, much more in line with those observed with $Z_{\mathrm{Te;\,SII}}$ and $Z_{\mathrm{Te;\,Y20}}$.

From diagnostic diagrams, we identify the presence of non-star-forming emission in the SAMI609396 system. We attribute this emission to shock-heated gas on the basis of the observed correlation between the [S II]/Hα emission line ratio and the measured velocity dispersion. Despite applying our analysis to the star-forming selected region around SAMI609396B, we note that in reality each spaxel will contain some amount of contaminating, non-star-forming emission. In particular, we show that spaxels in this star-forming selected region appear to form the beginning of a spatial mixing sequence,

indicating spatial variations in the fractional contribution of non-star-forming emission to the total emission. Given the different ways metallicities from different diagnostics can be affected by contaminating emission, these variations could help to explain differences in the apparent metallicity trends recovered.

Aside from the $Z_{\mathrm{Te;\,LS12}}$ application of the direct method, our metallicity measurements favour a flat metallicity gradient for SAMI609396B. This flat gradient can be explained by the effects of the merger which are known to produce flattened metallicity gradients due to inflow of pristine gas from large radii (Kewley et al., 2010). However, possible contamination from shock emission may affect the gradient measurement.

The direct method remains the main calibration baseline for studying the chemical evolution of galaxies. However, it is not immune to modelling uncertainties. This study highlights the importance of adequately constraining the internal ionisation and temperature structure within H II regions when probing spatial variations of the metallicity across galaxies. We have shown here that abundance measurements based on T_e([O III]) alone are not a good indicator of the metallicity gradient in SAMI609396B due to their sensitivity to the ionisation parameter.

Spatially resolved applications of the direct method are currently limited even within the local Universe. Low-mass galaxies ($<10^{9.5}\ M_{\odot}$) contribute significantly to the stellar mass density and escape fraction of hydrogen ionizing photons at high redshift. However, the internal chemical distribution of these low-mass galaxies are rarely constrained owing to the spatial resolution and detection limit. This situation will be improved by forthcoming facilities such as *JWST*/NIRSpec and ground-based ELTs, which will push both the depth and spatial resolution attainable for IFU observations. The in-depth analysis of local objects like SAMI609396B, thus sets the stage for future detailed metallicity analysis of low-mass galaxies at high redshift.

Chapter 5

Mapping extra-planar electron temperature variations in MRK 1486

5.1 Context

Galactic winds, driven either by supernovae (SNe) and stellar winds, or by active galactic nuclei (AGN), are a critical mechanism for regulating star formation and distributing matter in galaxies (Veilleux et al., 2005; Weiner et al., 2009; Hopkins et al., 2012; Rupke, 2018). The imprint of galactic winds can be observed indirectly due to their role in shaping galaxy scaling relations including the Mass-Metallicity Relation (Tremonti et al., 2004; Sanders et al., 2020a) and the stellar mass function of galaxies (Benson et al., 2003; Baldry et al., 2012). Additionally, absorption measurements of metals in the circumgalactic medium (CGM) suggest enrichment via outflows from the interstellar medium (ISM) (Tumlinson et al., 2011; Kacprzak et al., 2012; Werk et al., 2013; Borthakur et al., 2016). However, beyond these indirect probes, direct measurements of outflowing gas are required to further our understanding of the nature of galactic winds and the feedback processes that power them.

Historically, direct measurements of outflows are conducted either with blue-shifted rest-UV absorption lines (Heckman et al., 2000; Weiner et al., 2009; Chisholm et al., 2018), or with extra-planar emission lines (e.g. Heckman et al., 1990; Westmoquette et al., 2011). More recently, with the advent of integral-field unit (IFU) spectroscopy, extra-planar emission in outflowing galaxies can be mapped spatially, either for face-on or edge-on systems (Sharp & Bland-Hawthorn, 2010; Rich et al., 2010; Ho et al., 2014,

2016). Chisholm et al. (2018) used UV absorption lines from a sample of seven galaxies to determine that warm outflowing gas can be over ten times more metal rich than the ISM. However, extra-planar metallicity measurements have not been made with emission lines due to the associated systematic uncertainties.

Temperature is a fundamental property in understanding the nature of ionised gas. The temperature of extra-planar gas in an outflowing galaxy will depend on the balance of heating from the feedback source driving the outflows (e.g. SNe or stellar winds), and radiative cooling of the gas which is dependent on metallicity. However, observational measurements of the electron temperature of ionised gas depend on faint auroral emission lines that are extremely challenging to obtain. Attempts to map electron temperatures have therefore been limited, even in large nearby galaxies (e.g Berg et al. 2013, 2020; Ho et al. 2019; Leung et al. 2020 and work presented in Chapter 4) and direct electron temperature measurements in extra-planar regions have remained elusive, despite the clear utility they present.

Here we report the first example of direct electron temperature measurements of extra-planar gas using deep *Keck*/KCWI observations of the edge-on starbursting galaxy MRK 1486 which exhibits strong star-formation driven galactic winds. Our observations yield measurements of the [O III] λ4363 auroral line in individual spaxels with signal-to-noise up to $S/N \sim 40$ in the brightest regions. Spaxels with $S/N_{\lambda4363} > 3$ are observed up to 2.7 arsec (∼1.9 kpc) off the plane of the disk allowing the electron temperature to be mapped with unprecedented detail and spatial coverage in this extra-planar regime. MRK 1486 has a high star-formation rate (SFR $= 3.6 \pm 0.7$ $M_\odot$ yr^{-1}) relative to its mass ($\log(M_*/M_\odot) = 9.3 \pm 0.2$) placing it well above the star-forming main sequence. These parameters suggest MRK 1486 may be analogous to galaxies at high-redshift, an epoch where galactic winds were ubiquitous (Weiner et al., 2009). These data can afford unique constraints on the feedback sources powering winds in starbursting galaxies, as well as the nature of the extra-planer gas itself.

5.2 Data

MRK 1486 is a star-forming galaxy at $z = 0.033841$ with an r-band magnitude of $m_r = 16.81$. MRK 1486 has stellar mass $\log(M_*/M_\odot) = 9.3 \pm 0.2$ and star-formation

rate SFR $= 3.6 \pm 0.7\ M_{\odot}\ \mathrm{yr}^{-1}$ (Chisholm et al., 2018), placing it well above the star-forming main-sequence (SFMS) of typical local galaxies. Visually, it can be seen to be a disk galaxy oriented edge-on and has bursty morphology, presumably related to the high SFR (see Figure 5.1).

IFU observations of MRK 1486 were carried out on 2020 March 22nd with *Keck*/KCWI as part of the Deep near-UV observations of Entrained gas in Turbulent galaxies (DUVET) survey (Fisher et al., *in prep.*). The large IFU slicer setting was used which gives a spatial sampling of $0\farcs29 \times 1\farcs35$ over a $20'' \times 33''$ field-of-view. The observations were carried out in good seeing conditions ($\sim 0\farcs7$ at 0.5 μm). The spatial resolution is thus limited by the atmospheric seeing in the perpendicular direction (where spatial sampling is only $0\farcs29$). We used the blue medium-dispersion grating (BM) in two configurations ("blue" grating angle: 40.507°; "red" grating angle: 44.540°) to provide spectral coverage from 3731 Å to 5284 Å with spectral resolution $R \sim 2000$, affording measurements of the key [O II] $\lambda\lambda$3726, 9 and [O III] λ5007 emission lines, as well as other lines within that spectral region. Continuous spectral coverage is obtained, with the red and blue grating settings in fact overlapping between 4408 Å and 4627 Å. The red grating setting received total integration of 2260 s across 9 exposures. This involved seven long exposures (6×300 s and 1×400 s) as well as 2×30 s exposures. The long exposures in this strategy were used to improve the efficiency of detecting faint lines (e.g. [O III] λ4363), and the short exposures allowed for flux measurements of [O III] λ5007 which was saturated in the long exposures. The [O III] λ4959/λ5007 ratio was used to ensure the saturation fix was appropriately applied since the [O III] λ4959 line did not saturate in the longer exposures. The blue grating setting simply involved 7×300 s exposures for 2100 s of total integration.

5.2.1 Data reduction

The data were reduced with the KCWI Data Extraction and Reduction Pipeline v1.1.0. Since [O III] λ5007 was saturated for some spaxels in the long exposures of the red grating setting, the [O III] λ5007 / [O III] λ4959 ratio (which should be constant) was used to identify which spaxels were saturating in the long exposures. Spaxels with saturated [O III] λ5007 would have spectral pixels in proximity to that line replaced by measurements from the shorter exposures. We found that the CCD had a non-linear response

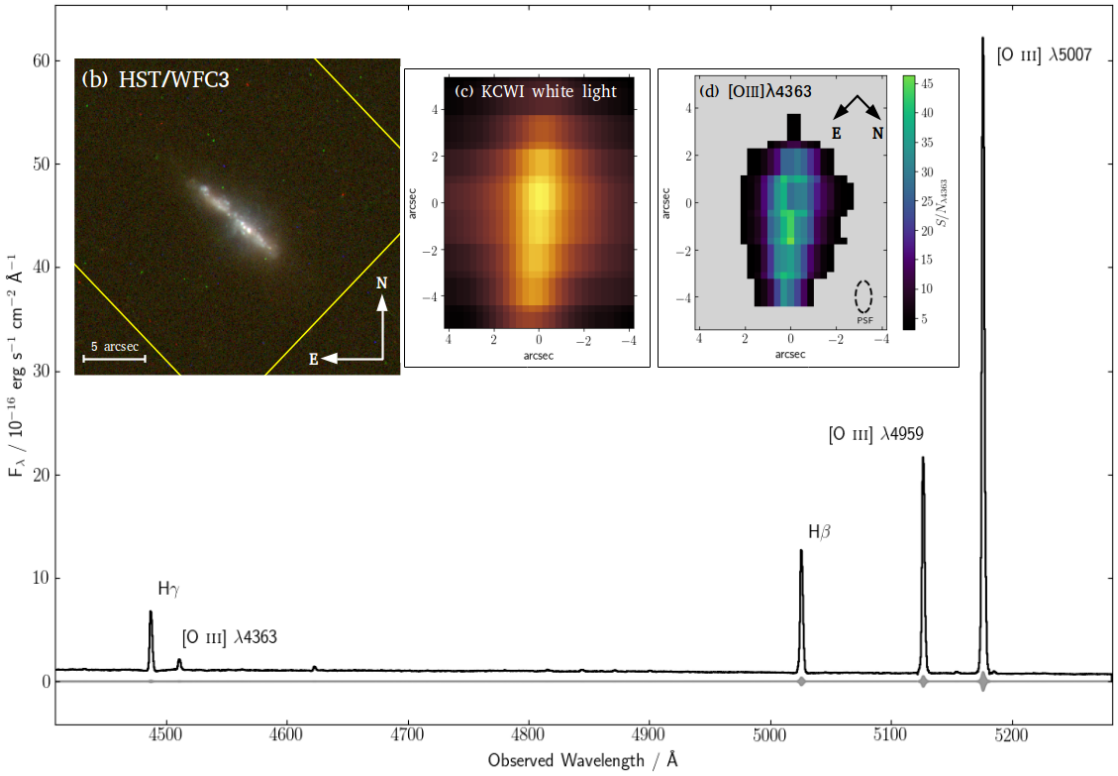

FIGURE 5.1: Overview of observational data for MRK 1486. *Main panel:* Red grating setting 1D spectrum from an example spaxel showing prominent emission lines. Spectrum is shown in black and 1-σ error spectrum is shown in grey. *Inset panel (b):* Color composite from *HST*/WFC3 broadband imaging (blue = F336W, green = F438W, red = F775W). The yellow box shows the extent of the KCWI field of view. *Inset panel (c):* Zoom-in of white light image from *Keck*/KCWI red grating setting shown with a logarithmic color scale with arbitrary normalisation. *Inset panel (d):* Signal-to-noise map of [O III] λ4363 shown for spaxels with [O III] $\lambda4363 \geq 3$. This defines the primary spatial region over which subsequent analysis is performed. The dashed ellipse shows the approximate PSF and applies to both panels (c) and (d).

and that this ratio would start to become lower than expected even before full saturation occured. Additionally, some adjacent spaxels which did not saturate had deviations from the expected ratio, suggesting that the saturation can "bleed" spatially. However, this was easily resolved with [O III] λ5007 replacements in these spaxels as well. Sky subtraction was performed within the pipeline using an off-object sky frame immediately before or after a set of exposures. As a final step the data cubes were re-sampled onto a regular $0\farcs29 \times 0\farcs29$ grid. Visual inspection of the reduced datacubes clearly indicate prominent detections in individual spaxels of the [O III] λ4363 auroral emission line which is critical to the temperature measurements described in this chapter.

5.3 Methods

5.3.1 Continuum subtraction

We fit the stellar continuum in each spaxel of the red and blue grating settings using `pPXF` (Cappellari, 2017). The observational data are rebinned to a logarithmic wavelength scale. Four-moment stellar continuum fits are performed to model spectra from the Binary Population and Spectral Synthesis code (BPASS, v2.2.1; Eldridge et al. 2017; Stanway & Eldridge 2018). We use spectra from the '135_300' IMF[1] which can be easily handled using the purpose built python package `hoki`[2] (Stevance et al., 2020). Two-moment fits are simultaneously performed on the Balmer emission lines and the strong [O II] $\lambda\lambda$3726, 9 and [O III] $\lambda\lambda$4959, 5007 collisionally excited lines, however these derived fluxes are not used in the final analysis. Since many faint emission lines are present in the MRK 1486 datacube, we first run a continuum fit on the summed global spectrum of MRK 1486 for each grating setting in which the iterative sigma-clipping approach of Cappellari et al. (2002)[3] is employed, to mask out these faint emission features. These masks are then applied to individual spaxels during the subsequent fitting routine to minimise the effect of faint emission features on the spectral template fitting. As a final step, we visually inspected spaxel fits to ensure that no spurious features were affecting the resulting fits. We then subtract the best-fit continuum from the observed data for each spaxel in preparation for emission line fitting.

5.3.2 Emission line fitting

MRK 1486 exhibits extended spatial emission out to at least 5″ off the axis of the disk (projected distance of 3.5 kpc). From visual inspection, the extended emission is clearly seen to exhibit complex kinematic behaviour. However, this study is focused on the [O III] λ4363 auroral emission feature which is only observed out to relatively smaller distances of 2.7″ ($\sim$1.9 kpc). In this region, the emission line profiles are well modelled by a single Gaussian component and the emission line fitting is conducted accordingly.

[1] Refer to Table 1 in Stanway & Eldridge (2018) for characterisation of this IMF. This is the default IMF used by Stanway & Eldridge (2018).

[2] https://github.com/HeloiseS/hoki

[3] This is contained within the `pPXF` package and can be used by setting the `clean=True` keyword in the main `pPXF` routine.

Emission line fits are performed on the continuum-subtracted data for each spaxel according to the following procedure: The velocity (or redshift, z) and velocity dispersion (σ) is obtained by simultaneously fitting single-component Gaussian profiles to the Hβ, Hγ, and [O III] $\lambda\lambda$4959, 5007, emission features in the red setting using a χ^2-minimisation procedure. From this we obtain median values of $z = 0.033841$ and $\sigma = 1.10$ Å ($FWHM = 2.60$ Å) across the data cube. Using the resulting kinematic values, which are treated as fixed for that spaxel, fluxes are fit individually for each emission line, using only a 30 Å sub-interval of the complete spectrum centred on the expected centroid of the line at the best-fit redshift. We fit the [Ne III] λ3869, Hδ, Hγ, and [O III] λ4363 lines in the blue grating setting, and the Hγ, [O III] λ4363, Hβ, [O III] λ4959, [O III] λ5007 lines in the red grating setting. Note that the [O III] λ4363 line is observable in both grating settings. Throughout this chapter, when the [O III] λ4363 / λ5007 ratio is used, only the [O III] λ4363 flux measurement from the red setting is considered, to mitigate any possible uncertainties arising from flux calibration. The fluxes of the [O II] $\lambda\lambda$3726, 9 doublet are fit simultaneously to each other since they are partially blended at this spectral resolution. No restrictions are imposed on the flux ratio of this doublet. The [O III] λ5007 / λ4959 ratio is often treated as being of fixed value. However, we fit the fluxes of these lines separately and use the value of this ratio to verify that the saturation of [O III] λ5007 has been appropriately corrected for in the data reduction (see Section 5.2.1). We find that 90 % of the spaxels fit have $2.83 \leq R_{\lambda5007/\lambda4959} \leq 3.09$, with deviations from this range being limited to the outermost regions where S/N is lowest.

As in emission line fitting described for Chapter 4, the uncertainty of each emission line is calculated by adding in quadrature the statistical uncertainty of the fit, and an estimate of the uncertainty in the continuum level obtained using a 30 Å box in close spectral proximity to the emission line which possesses no emission features (refer to Eq. 1 in Pérez-Montero 2017 for calculation). The analysis presented in this chapter is limited to spaxels where $S/N_{\lambda4363} \geq 3$. Hβ, [O II] $\lambda\lambda$3726, 9 and [O III] $\lambda\lambda$4959, 5007 emission lines used in this analysis are all detected easily with minimum $S/N > 10$ across this spatial region.

These 2D emission line maps are finally corrected for dust extinction based on the Hγ/Hβ flux ratio observed in the red grating setting. We assume an intrinsic Balmer decrement of $f_{H\gamma}/f_{H\beta} = 0.468$ (Dopita & Sutherland, 2003; Groves et al., 2012) and a Cardelli

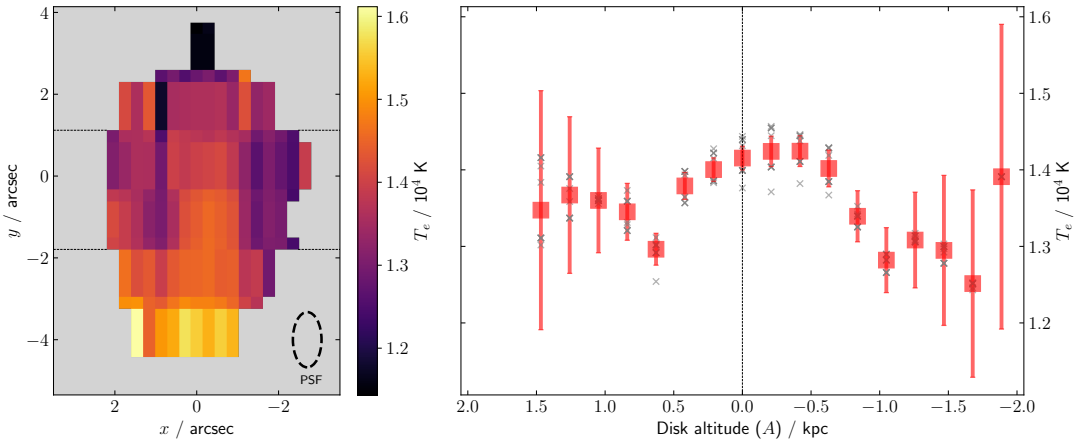

FIGURE 5.2: *Left:* Map of electron temperature as traced by [O III] λ4363 / λ5007 for the edge-on disk galaxy MRK 1486. The plane of the disk runs vertically in this map, centred on $x = 0$ arcsec. *Right:* Electron temperature profile as a function of physical distance perpendicular to the plane of the disk. Only spaxels in the region between $-1.79 \leq y \leq 1.12$ arcsec from the left panel are included in the right panel. Grey crosses indicate individual values, while the red points are median values with red error bars depicting the median uncertainty of values in that bin.

et al. (1989) extinction law with $R_V = 3.1$. Analysis presented in the remainder of this chapter is conducted using these reddening corrected emission line flux maps.

5.4 Electron temperature variations

We derive an electron temperature map for MRK 1486 from the [O III] λ4363 / λ5007 emission line ratio according to the relation set out by Nicholls et al. (2020)[4]. Derived temperatures fall in the range $1.15 \leq T_e/10^4 \text{ K} \leq 1.62$ as shown in the left panel of Figure 5.2. Looking vertically along the T_e map shown in the the left-hand panel of Figure 5.2, a trend of increasing T_e with increasing disk radius can be seen from the centre at $x = 0$ arcsec and $y = 0$ arcsec (where the white light flux peaks) along the disk down to $y \approx -4$ arcsec. This is likely the result of a negative radial metallicity gradient along the MRK 1486 disk, similar to those widely observed in low-redshift galaxies. This analysis will instead focus on extra-planar electron temperature variations.

[4]This mirrors the methodology employed earlier in Chapter 4.3.2 and the reader is referred to discussion in that section for more details.

5.4.1 Extra-planar variations

The right-hand panel of Figure 5.2 shows the temperature profile for the core region (bounded by the black dotted lines in the left panel) moving perpendicular to the plane of the disk of MRK 1486. The grey crosses show measurement values from individual spaxels while the red points depict median values in each "disk altitude" (A) bin. The red error bars show the median measurement uncertainty for each bin which generally represent the uncertainties of all values in that bin well. From this profile, we see a clear decrease in temperature moving away from the plane of the disk, particularly in the right-hand direction.

The electron temperature of the gas is determined by the balance of heating from the ionising source and radiative cooling. Emission from metals is the primary cooling mechanism in ionised gas. Accordingly, temperature is strongly dependent on metallicity with higher metallicity regions typically exhibiting lower temperatures. Indeed, if one were to simply input the T_e map from Figure 5.2 into a direct method metallicity calibration, a clear increase in metallicity moving away from the plane of the disk would arise. However, while an increase in metallicity will drive a decrease in temperature, *lower measured temperatures do not necessarily imply higher metallicities.* Indeed, in this sort of extra-planar study, it is reasonable to expect that emission at larger disk altitudes would be powered by a systematically different ionising flux compared to gas in the plane of the disk. Thus, it is paramount that a thorough analysis of extra-planar temperature variations must consider variations in the ionisation parameter.

Figure 5.3 shows a map of the $O32$ = log([O III] λ5007 / [O II] $\lambda\lambda$3726, 9) emission line ratio which is a widely used proxy for ionisation parameter. It can be clearly seen from the right-hand panel of Figure 5.3 that the $O32$ ratio (and hence likely the ionisation parameter) is decreasing with increasing disk altitude in both directions. This decrease in ionisation parameter may contribute to the observed decrease in electron temperature.

To explore the relative contribution of photoionisation heating and radiative cooling to the observed temperature variations, we compare observed [O III] λ4363 / λ5007 and $O32$ line ratios from MRK 1486 to predictions from MAPPINGS model grids (A. D. Thomas; private communication). These models were computed with MAPPINGS V (Sutherland & Dopita, 2017) for a STARBURST99 continuous star-formation model at

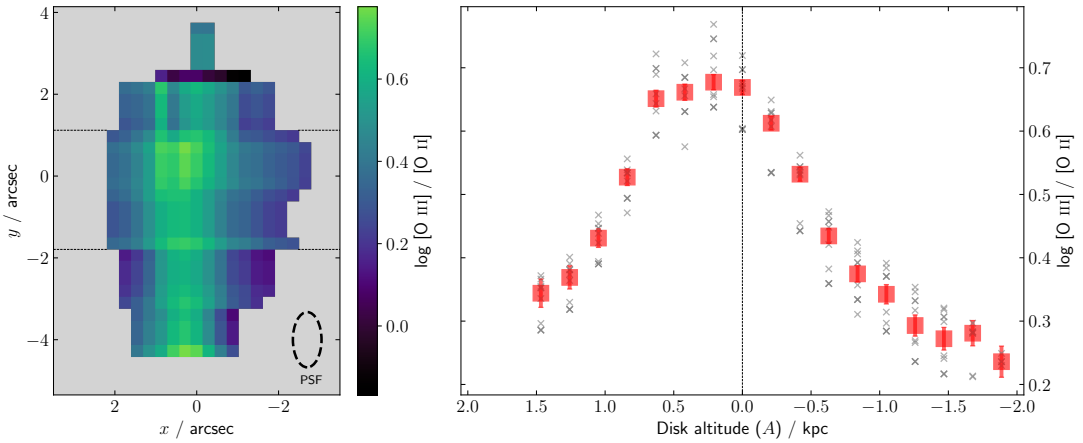

FIGURE 5.3: As for Figure 5.2 except showing $O32$ = log [O III] λ5007 / [O II] $\lambda\lambda$3726, 9 which is a widely used proxy for ionisation parameter. Here a $O32$ gradient can be seen, even more clearly than for the T_e. This suggests as expected that the ionisation parameter decreases moving away from the plane of the disk. This will contribute to the observed temperature variation.

$1 M_\odot$ yr^{-1} (Leitherer et al., 2014) over a range of metallicity, ionisation parameter, and ISM pressure values.

Figure 5.4 shows a subset of these model grids as line-connected colour-filled points. The grids shown depict three different metallicities, across a range of ionisation parameter values indicated by the colour coding.[5] As expected, the ionisation parameter correlates well with the $O32$ line ratio, while metallicity correlates negatively with [O III] λ4363 / λ5007 which is a proxy for electron temperature.

5.4.1.1 Contribution of ionisation parameter

Figure 5.4 shows the importance of considering the effects of ionisation parameter when interpreting electron temperature variations. MRK 1486 falls in a regime of low-metallicity below $12+\log(O/H) \lesssim 8.1$ where the MAPPINGS grids predict that a decrease in ionisation parameter will lead to a decrease in [O III] temperature. Of course, from Figure 5.3 we know that ionisation parameter is decreasing with increasing disk altitude, so the ionisation parameter will likely be a key contributor to the observed T_e variations in MRK 1486.

[5] All values shown are for ISM pressure of $\log(P/k) = 5.8$, and no significant variation from the grids shown was observed for other plausible pressure values.

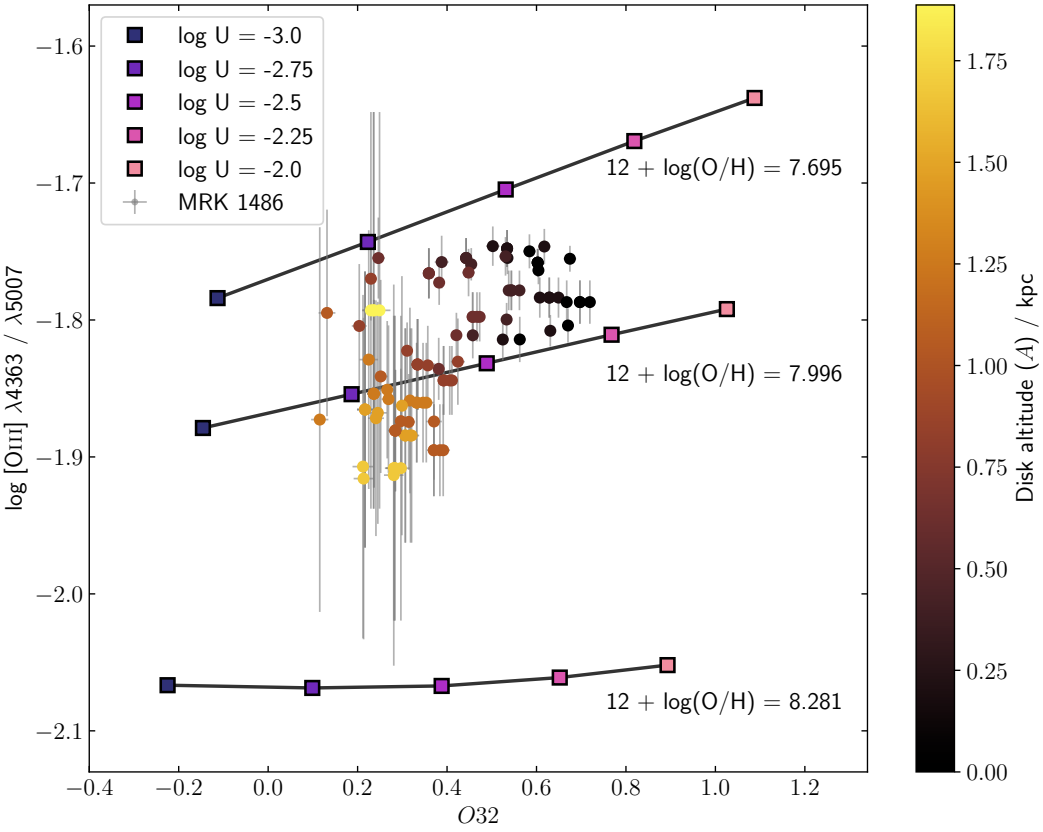

FIGURE 5.4: Evolution of line ratios with disk altitude in MRK 1486 as compared to MAPPINGS model predictions. The y-axis shows [O III] λ4363/λ5007 which correlates with T_e. The x-axis shows $O32$ = [O III] λ5007 / [O II] $\lambda\lambda$3726,9 which correlates with ionisation parameter (log U). Decreasing ionisation parameter contributes to the observed decrease in T_e moving away from the plane of the disk in MRK 1486, however there is tentative evidence for contribution from an increase in metallicity. Coloured circles show values for individual spaxels in MRK 1486 in the region bounded by $-1.79 \leq y \leq 1.12$ arcsec and $x \geq 0.0$ arcsec as mapped in Figure 5.2. Thus, these points represent variations in the right-hand direction of the two panels in Figure 5.2. The colour bar indicates the physical distance of these points from the plane of the disk. Large colour-filled squares depict predicted emission line ratios from MAPPINGS. Different ionisation parameter (log U) values are indicated by the colour coding in the legend. Solid black lines connect model values of equal metallicity. The values shown assume a constant ISM pressure (log(P/k) = 5.8). Different pressure values within the plausible range change the model predictions minimally.

Indeed, the MRK 1486 data points form something of a loose sequence from higher temperature, higher $O32$ points (upper-right in Figure 5.4) generally being those close to the plane of the disk ($A \sim 0$ kpc in colour bar), while the lower temperature, lower $O32$ points (lower-left in Figure 5.4) are those at large disk altitude ($A \sim 1.75$ kpc in colour bar; right-hand side of T_e map in Figure 5.2). This sequence qualitatively follows the model grid lines of constant metallicity, suggesting that varying ionisation parameter is indeed a key contributor to the extra-planar temperature variations observed in Figure 5.2.

5.4.1.2 Contribution of metallicity

In addition to the contribution from ionisation parameter, the diagnostic diagram in Figure 5.4 can test for the presence of metallicity variations in a simple way: if the gradient of the sequence of observational points is steeper than the gradient of the lines of constant metallicity from the model grids, this would suggest that metallicity is increasing moving away from the plane of the disk.

Considering the cluster of points above the $12 + \log(O/H) = 7.996$ model line with $\log U \geq -2.5$ and $A \sim 0$ kpc, and the cluster of $A \sim 1.75$ kpc points below this line at lower $\log U$, this could be considered as tentative evidence for an increase in metallicity moving away from the plane of the disk. Of course, the large uncertainties on the [O III] λ4363 / λ5007 line ratio at large distance makes it hard to ascertain the precise trend. Furthermore, ionising spectra from different stellar population models may alter the $\log(O/H)$ – $\log U$ slope in these photoionisation model grids which could affect the conclusions drawn around whether metallicity variations are present.

It has been shown that galactic outflows can be metal-enriched compared to the average ISM abundance (e.g. Chisholm et al., 2018). Thus, a metallicity increase moving away from the disk would not be unexpected as it could reflect the imprint of these chemically enriched outflows. However, we defer drawing extensive conclusions until further analysis has been conducted. Future work for this project will focus on expanding the photoionisation modelling analysis to include different stellar population models to rigorously evaluate the possible presence of extra-planar metallicity variations.

5.5 Summary

In this chapter we have presented preliminary results outlining the first spatially resolved direct electron temperature measurements of extra-planar gas in an outflowing galaxy. Deep IFU spectroscopy from *Keck*/KCWI yielded $S/N > 3$ detections of the [O III] λ4363 auroral emission line at angular separations of up to $2\farcs7$ beyond the plane of the disk in individual $0\farcs29$ resolution elements, corresponding to physical distances of almost 2 kpc off the plane of the disk.

A clear trend of decreasing electron temperature moving away from the plane of the disk was observed. Similarly, a trend of decreasing ionisation parameter with increasing disk altitude was observed, with early results from photoionisation modelling suggesting that this is a key contributor to the observed decrease in temperature. Lower electron temperature is often associated with higher metallicities and early results from photoionisation modelling suggest there is tentative evidence for increasing metallicity moving away from the plane of the disk. This would be in agreement with recent observations of metal-enriched outflows from UV absorption lines (Chisholm et al., 2018), however further analysis is required to rule out the effect that different stellar population models may have on this finding. We also see evidence of a radial temperature gradient, likely the result of an expected radial metallicity gradient.

Understanding the interplay of star-formation fuelled by infalling gas, and powerful feedback-driven outflows is critical to developing models of galaxy assembly, particularly during the epoch of peak cosmic star-formation (i.e $z \sim 1 - 3$). The spatially resolved electron temperature measurements presented here for MRK 1486 provide a novel probe for understanding this "galactic fountain" model. These observations will provide unique constraints on the feedback sources that drive galactic winds, and the effects that gas flows have on shaping the evolution of galaxies.

Chapter 6

Conclusions and future prospects

Studies of the chemical evolution of galaxies provide powerful insights into galaxy assembly histories. Gas is enriched with metals by nucleosynthesis from evolved stars. However, gas-phase metallicity is sensitive not only to star-formation histories, but also to galaxy-scale gas flows, and thus probes the effects of feedback physics, gas accretion, interactions and mergers. This array of processes probes not only the complex baryonic physics occurring within galaxies, but relies critically on their host dark matter haloes as these are the dominant contributor to galaxy gravitational potential wells.

The goal of this thesis has been to explore observational methods for constraining the chemical evolution of galaxies, with a particular emphasis on the performance of these methods near the capability limits of observational facilities. A primary focus has been the use of emission lines to measure chemical abundances in galaxies, and in particular better evaluate the systematic uncertainties involved in these methods, which have long been poorly constrained. This thesis has also explored measuring large-scale clustering of galaxies with space-based imaging surveys. Such measurements can in turn place constraints on the typical host dark matter halo masses of galaxy populations. This chapter provides a summary of the work presented in this thesis.

6.1 Clustering measurements of high-redshift galaxies

Galaxy clustering measurements are a powerful technique for drawing statistical connections between galaxy populations and the dark matter haloes that host them. While

galaxy clustering can be effectively characterised out to redshifts of $z \sim 2$, traditional correlation function measurements are difficult to apply at the highest redshifts due to the effects of cosmic variance and the limited survey volumes typically available.

Chapter 2 tested the efficacy of a statistical counts-in-cell approach to the measurement of large-scale galaxy bias for $z \sim 2$ galaxies from the Brightest of Reionising Galaxies (BoRG) survey, a pure-parallel survey conducted with *HST*/WFC3. This work presented the application of two different implementations: those outlined by Robertson (2010) and López-Sanjuan et al. (2015). We derived number counts of photometric $z \sim 2$ candidates across the 141 uncorrelated fields from the BoRG survey in order to undertake the first application of these methods to a space-based pure-parallel survey.

We found that for our sample, with a mean number count of $N = 6.9$ candidates per field, applying these methods yielded large-scale galaxy bias measurements broadly consistent with high-redshift bias measurements reported in the literature from correlation function measurements. We found the Robertson (2010) method to be more robust toward lower number counts where Poisson noise becomes increasingly significant, and ultimately suggest that mean number counts of $N \gtrsim 5$ per field are required for meaningful measurements to be made.

6.1.1 Future prospects: galaxy bias at $z \sim 8$ with JWST

Predictions for *JWST*/NIRCam suggest that an imaging survey of similar design to BoRG could comfortably detect dozens of galaxies per field at $z \sim 6$ and mean number counts at $z \sim 8$ of $N \sim 7$ per field, comparable to those presented here at $z \sim 2$ with BoRG. This suggests that application of these methods to such a survey could provide novel measurements of the large-scale clustering of galaxies in the epoch of reionisation. These clustering measurements would in turn enable unique constraints to be placed on the host dark matter haloes of galaxies in these early epochs of galaxy evolution.

6.2 Metallicity measurements with emission lines

6.2.1 Extending high-redshift metallicity measurements to low-mass

Galactic outflows will have the most significant effect in low-mass galaxies, motivating the desire to extend high-redshift Mass-Metallicity Relation (MZR) measurements to dwarf galaxies. However measuring emission lines in the near-infrared for faint galaxies at large cosmic distances presents a significant observational challenge.

Chapter 3 presented a case study for extending the $z \sim 1-2$ MZR to masses below $M_* = 10^9\ M_\odot$, examining in detail metallicity measurements made for a gravitationally lensed dwarf galaxy at $z = 1.00$ with $\log(M_*/M_\odot) = 8.62^{+0.10}_{-0.06}$. Using emission line measurements of Hα, [N II] λ6583 and [S II] $\lambda\lambda$6716, 6731 made from J-band spectroscopy from *Keck*/MOSFIRE, we derived metallicity from two different strong line diagnostics: $N2$ which depends only on [N II]/Hα, and $N2S2H\alpha$ which additionally incorporates information from the [N II]/[S II] line ratio. In particular, we used photoionisation model grids to explore the systematic uncertainty induced by possible variations in the ionisation parameter. We found that while the $N2$ diagnostic had a statistical uncertainty of only 0.05 dex, considering the effect of an unconstrained ionisation parameter leads to a much larger total uncertainty of 0.5 dex. However, while the $N2S2H\alpha$ metallicity suffered from a large statistical uncertainty (0.13 dex) due to modest constraints on the [N II]/[S II] line ratio in this detection limit regime, its improved handling of ionisation parameter degeneracy lead to an overall improvement in the total uncertainty (0.36 dex).

This suggests that the primary concern, as far as placing tight constraints on metallicity in high-redshift dwarf galaxies, is not the reduced signal-to-noise of emission line measurements. Rather it is the reduced *number* of emission lines available. Metallicity measurements made with simple diagnostics, such as $N2$, may do a poor job of constraining the nebular conditions and provide misleading metallicity measurements. This underlines the importance of assembling large suites of emission lines when exploring metallicities in high-redshift galaxies. The ionising conditions in high-redshift galaxies may be fundamentally different to local galaxies. Studying the chemical evolution of galaxies across cosmic time therefore requires that systematic evolution in the nebular conditions of galaxies can be adequately accounted for.

6.2.1.1 Future prospects: metallicity at high-redshift with JWST

High-redshift galaxy evolution studies will be revolutionised by the improved sensitivity and near-infrared spectral coverage of the Near Infrared Spectrograph (NIRSpec) on the forthcoming James Webb Space Telescope. In particular, by being above the atmosphere, *JWST*/NIRSpec will have uninterrupted spectral coverage across its $0.6 - 5$ micron wavelength range, vastly improving our improving our ability to assemble large suites of emission lines for high-redshift galaxies. These data sets will be critical in developing a more complete understanding of the nebular and ionising conditions in galaxies within the first $\sim$1 Gyr after the Big Bang. This in turn will be critical to understanding chemical evolution in the earliest epochs of galaxy evolution.

6.2.2 Spatially resolved auroral lines as a probe of chemical evolution

Faint auroral emission lines have long been coveted for the wealth of information available from the electron temperature (T_e) measurements they afford. In particular, T_e measurements can place vital constraints on chemical abundances. Historically, spatially resolved T_e studies have been scarce even in the nearby Universe due to the associated observational challenge. However, with advances in modern IFU spectroscopy this picture is changing, motivating a desire for more detailed integral-field auroral line studies.

6.2.2.1 Spatially resolved auroral lines in the SAMI Galaxy Survey

Chapter 4 presented the detailed study of SAMI609396B, a galaxy identified in archival data from the SAMI Galaxy Survey Public Data Release 2 as having prominent spatially resolvable emission of three auroral lines: [O III] λ4363, [S II] $\lambda\lambda$4069,76, and [S III] λ6312. SAMI609396B is undergoing a merger and has a high star-formation rate (4.21 ± 0.30 $M_\odot$ yr^{-1}) relative to its modest stellar mass ($\log(M_*/M_\odot) = 9.18 \pm 0.05$). These properties are quite distinct from the large spiral galaxies in which previous spatially resolved electron temperature measurements have been made. These conditions might instead be more analogous to the conditions in high-redshift galaxies ($z \sim 1$).

Our auroral line measurements afforded spatially resolved electron temperature measurements for two temperature zones: T_e([O III]) and T_e([S II]). Despite what is often

assumed, we found no evidence suggesting a positive correlation between these two temperature measurements across the spatial extent of SAMI609396B.

Utilising our T_e measurements, we derived spatially resolved "direct method" metallicity measurements. Critically, we found that the derived metallicity gradient of SAMI609396B varied dramatically based on our treatment of the temperature structure of the emitting H II regions. Adopting a default assumption of a fixed positive correlation between the temperature of the [O III] and [O II] zones yielded a strong positive gradient. However, when we accounted for possible deviations from this relation, either by applying an empirical correction based on the [O III]/[O II] strong line ratio (as recently outlined by Yates et al., 2020), or by deriving T_e([O II]) instead from our T_e([S II]) measurement, we obtained a flattened gradient, much more in line with the expectation for an interacting galaxy. Strong-line measurements made with several diagnostics also favoured a flattened gradient in SAMI609396B.

The ionising flux has an important contribution to the temperature structure within H II regions. These findings clearly show that variations in the ionisation parameter can masquerade as variations in the metallicity if naive assumptions are made about the temperature structure of H II regions. Direct method metallicity measurements made with auroral lines will continue to be an important technique in chemical evolution studies. However, this study has highlighted the importance of, where possible, obtaining auroral emission from multiple ionic species such that variations in the temperature structure can be accounted for.

6.2.2.2 Probing the conditions of extra-planar gas with auroral lines

Chapter 5 presented some preliminary findings of a study probing extra-planar temperature variations in MRK 1486, an edge-on starbursting galaxy exhibiting prominent outflows. These temperature measurements were afforded by strong detections of the [O III] λ4363 auroral line in individual $0\farcs29$ ($\sim$200 pc) resolution elements out to distances of almost 2 kpc off the plane of the disk.

Derived electron temperature measurements show clear trend of decreasing T_e with increasing disk altitude. Additionally, we find a strong decreasing trend in ionisation parameter moving away from the disk. Preliminary results from photoionisation modelling

show that, given the low metallicity of MRK 1486 ($12+\log(O/H) \lesssim 8.0$), this decrease in ionisation parameter likely provides an important contribution to the observed decrease in electron temperature. This modelling also suggests tentative evidence that part of the decrease in temperature could be attributed to an *increase* in metallicity moving away from the plane of the disk, in line with expectations from metal-enriched outflows (e.g. Chisholm et al., 2018). However at this time, further analysis is required to rule out the impact of different stellar populations on this conclusion.

6.2.2.3 Future prospects: large samples of spatially resolved auroral line measurements

While spatially resolving auroral lines has historically been the domain of single object studies, there is evidence that this is changing. Already modern IFU instruments on 8-10 m class telescopes have proven capable of mapping auroral lines in low-redshift galaxies (*VLT*/MUSE: Ho et al. 2019; *Keck*/KCWI: Leung et al. 2020 and Chapter 5). Furthermore, over the next decade we will see 20-30 m class telescopes such as the Extremely Large Telescope (*ELT*) coming online, offering improved sensitivity and spatial resolution. This will allow auroral lines to be measured in a much larger array of galaxies, but may also see the transition of 8-10 m class telescopes to more of a workhorse role, where assembling larger samples of the sort of deep observations presented in Chapter 5 may become more feasible. Additionally, with the integral-field mode of the forthcoming *JWST*/NIRSpec, our IFU capabilities in the near-infrared will be drastically improved, perhaps affording the first spatially resolved auroral line measurements at $z \gtrsim 1$.

The two studies presented in Chapters 4 & 5 highlight the power auroral lines possess in probing the nebular conditions of galaxies in great detail. While both studies showed that variations in the ionisation parameter can have a significant impact on measured electron temperatures (and the metallicities derived from those), it bears remembering that systematic uncertainties in auroral line studies still represent a significant improvement when compare to metallicity studies performed with strong-lines. For example, the uncertainties associated with making extra-planar metallicity measurements with strong-line methods are simply too large. However, Chapter 5 has shown that auroral lines can begin to shed light in this area. As we continue strive to constrain how the cycle of inflows, star-formation, and feedback-driven galactic winds shape the assembly of

galaxies, spatially resolved auroral line studies will provide valuable insights that reach beyond what is possible with strong-lines alone.

6.3 Concluding remarks

This thesis is written at a very exciting time for chemical evolution studies. The unprecedented sensitivity and wavelength coverage of *JWST* for near-infrared studies will vastly expand our understanding of chemical evolution during the early epochs of galaxy assembly, while the improved spatial resolution afforded by the advent of ground-based 30 m class telescopes will dramatically improve our ability to probe the internal structure of galaxies.

This thesis has primarily focused on exploring the performance of observational methods at the fringes of what is possible with current instruments. A recurrent theme in these studies has been the importance of proper treatment of systematic uncertainties, particularly regarding the complexity involved with probing the conditions of H II regions with emission lines. While the next generation of observational facilities will make what was once extraordinary become routine, ultimately these considerations will only become more important as we push chemical evolution studies of galaxies to new frontiers.

Appendix A

Chapter 2 Appendix

A.1 Primary Targets of the Most Overdense Fields

The coordinates of each pointing in a pure-parallel survey such as the BoRG survey are determined by those of a primary observation. These primary observations will generally be drawn from a large variety of different programs resulting in a set of of pointings distributed randomly across the sky. This randomness is important for our clustering measurement as we require a sample of fields that is not systematically biased toward regions of high density.

However, specific primary observations targeting regions of high density have the potential to yield pointings with high number counts that could be considered contaminants to this analysis. An example of this would be a redshift $z \sim 2$ quasar. Such an object could be expected to reside in a high density environment and an observation taken within $\sim$6 arcmin of this could bias our sample.

We have summarised the primary observations of fields with scaled counts of 12 or higher in table A.1 to ensure our analysis is not significantly biased in this way. This cut-off corresponds to a probability of $P(n \geq 12) \approx 0.05$ according to the associated Poisson distribution. Two of these fields were imaged in parallel to a primary targeting either an exo-planetary system (ExoPl) or a galactic star. Galactic targets will have no impact in biasing this survey. Similarly, low redshift galaxies, such as the redshift $z = 0.167$ galaxy targeted in this set of fields, are not an issue. The majority had primary programs targeting quasars (QSO). However, all of these targets had redshifts of $z < 1$ (much lower

TABLE A.1: Summary of the primary targets of the most overdense fields from the survey. The *Obj. Type* column describes the type of object targeted by the primary observation (QSO is a quasar, ExoPl is an exo-planetary system, Star a galactic star, Gal a galaxy, and AbLS an absorption line system) while column z gives its redshift. The field name and its density of counts per 4.41 arcmin^2 are given for reference.

Field name	Scaled count	Primary target	
		Obj. Type	z
2229-0945	22.88	QSO	0.798
2351-4332	18.47	QSO	2.885
1558+0812	18.30	QSO	0.517
1416+1638	17.12	QSO	0.744
1510+1115	13.92	QSO	0.285
1358+4326	13.74	ExoPl	galactic
0456-2203	13.02	QSO	0.533
1051+3359	12.43	Gal	0.167
0314-6712	12.36	Star	galactic
1525+0955	12.34	AbLS	1.278
1154+4639	12.32	QSO	0.643
1459+7146	12.26	QSO	0.905
0914+2822	12.02	QSO	0.735

than our redshift range), with the exception of one field (2351-4332). In this case the quasar is at a redshift of $z = 2.885$, sufficiently higher than our range of interest to avoid any impact. The remaining target was an absorption line system (AbLS) outside our redshift range of interest ($z = 1.278$). Thus, we concluded that none of the most overdense fields was physically associated to the primary target of the HST observations.

A.2 Tabulated Field Counts

A full list of $z \sim 2$ candidate counts from the fields analysed is given in tables A.2 – A.5.

TABLE A.2: Table of $z \sim 2$ candidate counts from the BoRG[z8] data set. Raw candidate counts are given in the second column. Areas are given in arcmin^2. The final column shows count rescaled to correspond to an area of 4.41 arcmin^2

Field name	$z \sim 2$	Area arcmin^2	Scaled count
0110-0224	10	13.81	3.19
0214+1255	1	4.42	1.00
0228-4102	4	4.43	3.98
0240-1857	5	4.45	4.96
0436-5259	4	4.33	4.07
0439-5317	5	4.28	5.15
0440-5244	5	4.34	5.08
0456-2203	13	4.40	13.02
0540-6409	7	3.95	7.82
0553-6405	4	4.00	4.41
0624-6432	5	4.44	4.96
0624-6440	5	4.39	5.02
0637-7518	13	6.33	9.06
0751+2917	8	4.52	7.80
0756+3043	4	4.19	4.21
0808+3946	7	4.30	7.18
0819+4911	6	4.52	5.85
0820+2332	7	4.46	6.91
0835+2456	10	4.43	9.95
0846+7654	11	4.41	11.00
0906+0255	8	4.39	8.03
0909+0002	6	4.40	6.02
0914+2822	12	4.40	12.02
0922+4505	3	4.38	3.02
0926+4000	4	4.45	3.96
0926+4426	7	4.49	6.88
0951+3304	8	4.41	8.00
0952+5304	4	4.42	3.99
1010+3001	11	4.54	10.68
1014-0423	1	4.48	0.98
1031+3804	3	4.48	2.95
1031+5052	8	5.55	6.36
1033+5051	6	5.50	4.81
1051+3359	12	4.26	12.43
1059+0519	9	4.43	8.96
1103-2330	9	4.37	9.09
1111+5545	6	4.31	6.15
1118-1858	3	4.23	3.13
1119+4026	7	4.46	6.92
1131+3114	6	4.41	5.99
1152+5441	5	4.40	5.02
1153+0056	4	4.44	3.97
1209+4543	6	4.42	5.98

TABLE A.3: Continuation of Table A.2

Field name	$z \sim 2$	Area arcmin^2	Scaled count
1230+0750	6	4.24	6.24
1242+5716	10	4.29	10.29
1245+3356	10	4.43	9.95
1301+0000	1	4.44	0.99
1337+0028	6	4.47	5.92
1341+4123	7	4.36	7.07
1358+4326	14	4.49	13.74
1358+4334	8	4.32	8.16
1408+5503	4	4.32	4.08
1416+1638	17	4.38	17.12
1429-0331	8	4.35	8.10
1437+5043	9	6.53	6.08
1459+7146	12	4.32	12.26
1510+1115	14	4.43	13.92
1524+0954	6	4.32	6.12
1555+1108	7	4.31	7.16
1632+3733	2	4.37	2.02
1632+3737	7	4.37	7.06
2057-4412	8	4.34	8.13
2132+1004	3	4.46	2.96
2132-1202	8	4.37	8.08
2155-4411	10	4.49	9.82
2203+1851	8	4.60	7.66
2313-2243	3	5.59	2.37
2345+0054	2	4.48	1.97
2351-4332	18	4.30	18.47

TABLE A.4: Table of $z \sim 2$ candidate counts from the BoRG[z9] data set. Raw candidate counts are given in the second column. Areas are given in arcmin^2. The final column shows count rescaled to correspond to an area of 4.41 arcmin^2

Field name	$z \sim 2$	Area arcmin^2	Scaled count
0116+1425	6	4.29	6.17
0119-3411	0	4.07	0.00
0132-7326	10	4.31	10.23
0235-0357	4	5.93	2.97
0314-6712	13	4.64	12.36
0337-0507	5	4.50	4.89
0554-6005	11	4.16	11.67
0751+2917	9	4.46	8.91
0807+3606	3	4.46	2.96
0851+4240	9	4.37	9.07
0853+0310	3	4.45	2.97
0925+1360	7	4.44	6.95
0925+3439	4	4.43	3.98
0949+5759	5	4.26	5.18
0953+5150	8	4.42	7.98
0953+5153	8	4.58	7.71
0953+5157	5	4.43	4.98
0955+4528	2	4.48	1.97
0956+2848	5	4.43	4.98
1017-2052	9	4.37	9.08
1018+0544	8	4.40	8.02
1048+1518	10	4.40	10.02
1103+2913	7	4.41	7.00
1104+2813	7	4.37	7.06
1106+2855	4	4.43	3.98
1106+2925	7	4.46	6.93
1106+3508	6	4.39	6.02
1115+2548	7	4.49	6.87
1136+0747	6	4.42	5.98
1142+2640	4	4.50	3.92
1142+2647	8	4.42	7.98
1142+3020	5	4.45	4.96
1143+3019	9	4.40	9.03
1149+2202	1	4.40	1.00
1152+3402	3	4.31	3.07
1152+5434	13	5.69	10.08
1154+4639	12	4.30	12.32
1160+0015	10	4.44	9.94
1209+4543	6	4.40	6.01
1218+3008	7	4.31	7.16

TABLE A.5: Continuation of Table A.4

Field name	$z \sim 2$	Area arcmin^2	Scaled count
1259+4128	14	6.06	10.19
1334+3131	3	4.50	2.94
1410+2623	9	4.40	9.02
1413+0918	2	4.35	2.03
1431+0259	0	4.08	0.00
1437-0150	7	4.35	7.09
1438-0142	8	4.44	7.95
1442-0212	8	4.41	7.99
1503+3645	5	4.32	5.11
1519-0746	2	4.44	1.99
1520-2501	1	4.29	1.03
1524+0956	2	4.34	2.03
1525+0955	12	4.29	12.34
1525+0960	5	4.28	5.16
1536+1410	6	4.45	5.94
1558+0812	18	4.34	18.30
1607+1332	6	4.43	5.97
1614+4856	3	4.38	3.02
1619+2541	4	4.33	4.08
1632+3736	7	4.65	6.64
1659+3732	6	4.29	6.17
1708+4237	6	4.33	6.11
1715+0455	4	4.25	4.15
1715+0502	3	4.25	3.11
1738+1839	7	4.22	7.32
2008-6610	8	4.24	8.32
2057-1423	7	4.29	7.20
2134-0708	11	4.47	10.85
2140+0241	6	4.44	5.96
2228-0955	4	4.42	3.99
2229-0945	23	4.43	22.88
2253-1411	10	4.45	9.91
2312-1423	5	4.42	4.99
2323+0059	6	4.10	6.45

Appendix B

Chapter 4 Appendix

B.1 SAMI Galaxies with Auroral Lines

We provide in Table B.1 a list of galaxies in SAMI Data Release 2 public data showing visually identifiable [O III] λ4363 emission. We do not claim that this list is exhaustive; rather, it is intended to provide a starting point for any future work hoping to make use of auroral line detections in SAMI data. This list was compiled during an exploratory search of the SAMI Data Release public data by visually inspecting the 1D spectra obtained by binning spaxels with the highest signal-to-noise on the Hα emission line. All cases other than SAMI609396 required some degree of spatial binning to achieve $S/N_{\lambda 4363} \gtrsim 5$. In our brief exploration we found that typically fewer than $\sim$4-5 usable bins could be extracted, however we did not expend any effort optimising these binning schemes. Our search focused on the [O III] λ4363 auroral line, however we note that in many of the galaxies listed in Table B.1, [S II] $\lambda\lambda$4069, 76 and [S III] λ6312 are also clearly

Table B.1: SAMI Galaxies with visually identifiable [O III] λ4363 emission.

SAMI ID
84107
137071
177518
209319
325376
561143
567676
567736
609396

present. We speculate that there may be galaxies with prominent [S II] and [S III] auroral line emission which were not picked up in our [O III] based search. We note that the [N II] λ5755 and [O II] $\lambda\lambda$7320, 30 auroral lines are typically not observable with SAMI. The [N II] λ5755 auroral line falls in the wavelength gap between the blue and red arms in the SAMI datacubes. The [O II] $\lambda\lambda$7320, 30 line falls near the red limit of the SAMI Galaxy Survey data and often outside the spectral coverage. Even in cases where it falls inside the spectral coverage, we find it is not detectable.

B.2 Global properties

SAMI DR2 value added data products include a spatially resolved star-formation rate (SFR) map based on measured Hα flux (refer to Medling et al. 2018 for details). We derive a global SFR for SAMI609396B by summing the spaxel by spaxel star-formation rate over the SAMI609396B selection mask defined in Section 4.2.2, obtaining SFR = $4.21 \pm 0.30 M_{\odot}$ yr^{-1}. However, it is worth noting that the spatial region considered here is limited by the SAMI field-of-view which does not achieve full coverage of SAMI609396B (see Fig 4.1). Indeed, the star-formation rate map (Fig 4.2 panel c) appears to peak near the FoV boundary. It is likely that the region extending beyond the FoV contributes significantly to the global SFR of SAMI609396B. In that sense, we suggest that the quoted SFR can be considered as a lower-bound. In addition to SAMI609396B, we derive SFR for the more massive companion galaxy by summing the SFR map over the remainder of the SAMI FoV. This yields a value of SFR = $0.32 \pm 0.08 M_{\odot}$ yr^{-1}, although we note that this spatial region exhibits significant contribution to its emission spectrum from non-star-forming sources which may bias this value (§ 4.5.2).

TABLE B.2: Deblended magnitudes used to estimate stellar masses for SAMI609396B and companion

	g-band	i-band
SAMI609396B	15.017 ± 0.001	14.932 ± 0.001
Companion	15.064 ± 0.001	14.351 ± 0.001

We derive global stellar mass values from g- and i-band photometry using the relation described in Section 4.2 of Bryant et al. (2015), based on stellar mass estimates from Taylor et al. (2011). We create a deblended segmentation map for the SDSS g-band imaging using `detect_sources` and `deblend_sources` from `photutils` package (Bradley et al.,

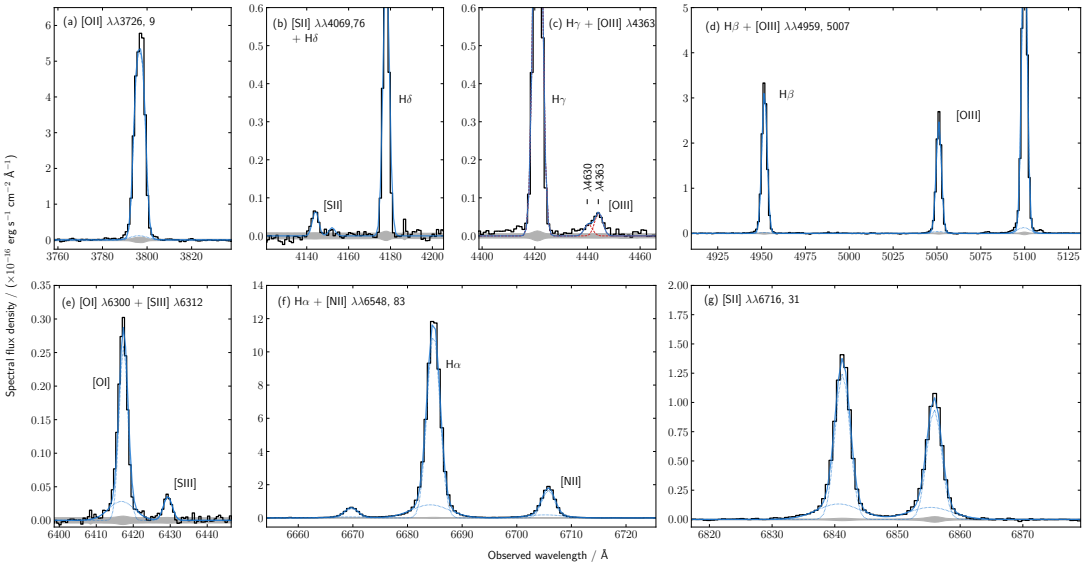

FIGURE B.1: Emission line fits to a 1D spectrum from an individual spaxel. Each panel (a) – (g) highlights different emission lines used in this analysis as described in the inset text. Units on each axis are the same for each panel, however note the normalisation on the vertical axis varies according to the strength of the emission lines shown. In each panel the continuum subtracted spectrum is shown as the black step plot, while the grey shaded band shows the 1-σ error spectrum. The best-fit emission line model is shown by the solid blue line, while the blue dotted lines show the individual broad- and narrow-components fit to the profile, where applicable. The three auroral lines observed in SAMI609396B, [SII]$\lambda\lambda$4069, 76, [OIII]λ4363, and [SIII]λ6312, are shown in panels (b), (c), and (e) respectively. The red dashed lines in panel (c) show the individual profiles of the [FeII]λ4360 and [OIII]λ4363 emission lines, fit simultaneously to account for blending of these features (see §B.3.3).

2019). We use default values of 32 multi-thresholding levels and a blending contrast of 0.001 to run the deblending. The magnitudes obtained from these segmentation images are given in Table B.2. Applying the Bryant et al. (2015) relation to those magnitudes we obtain stellar mass estimates of $\log(M_*/M_\odot) = 9.11 \pm 0.10$ for SAMI609396B and $\log(M_*/M_\odot) = 9.78 \pm 0.10$ for its more massive companion. The 0.10 dex uncertainties reflect the quoted 1-σ scatter in this relation (Taylor et al., 2011). We ignore the flux uncertainties from the SDSS imaging as these contribute only 0.001 dex variations in stellar mass. These values correspond to to a mass ratio of $\sim$0.21 for this merger system The global SFR and M_* values derived here place SAMI609396B at least 1.3 dex above the star-forming main sequence (SFMS) for local star-forming galaxies (Renzini & Peng, 2015).

B.3 Spectral Fitting

The SAMI DR2 value-added release includes emission line maps for the most widely used strong optical lines. However, the primary focus of this work is the auroral emission lines, which are not included in these data products. Here we outline the methods behind our own spectral fitting to the SAMI DR2 data cubes. Flux measurements obtained here are used throughout § 4.3-4.4.

B.3.1 Continuum Subtraction

For each spaxel in the reduced SAMI datacube, we fit the stellar continuum of the blue and red arms simultaneously using `pPXF` (Cappellari, 2017). The 1D spectra for each spaxel are logarithmically rebinned. Given the different wavelength resolutions of the two arms, the higher resolution red arm is sampled down to match the velocity scale of the lower spectral resolution blue arm. Four moment fits are performed to the stellar continuum using the MILES library of stellar templates (Vazdekis et al., 2010). Two moment fits to the Balmer emission lines and strongest forbidden emission lines ([O III]$\lambda\lambda$4959,5007, [N II]$\lambda\lambda$6548,83 and [S II]$\lambda\lambda$6716,31) are included in the fitting procedure, however these derived emission line fluxes are discarded (refer to §B.3.2 for emission line fitting).

SAMI609396 exhibits strong emission lines with high equivalent-widths ($>$ 200 Å) and thus even spectra from individual spaxels feature many faint emission lines that are not widely studied. To ensure that the effect of these faint emission lines on the stellar template fitting is minimised, an iterative sigma-clipping method is employed (described in §2.1 of Cappellari et al. 2002). In this approach, once a global minimum is found, spectral pixels deviating more than 3σ from this best-fit template are masked out and a new global fit is obtained. This process is repeated until no additional pixels are masked out. As a final step, the fits were visually inspected to ensure no spurious spectral features affected the fitting.

A continuum subtracted spectrum was obtained for each spaxel by subtracting the best-fitting stellar spectrum from the reduced observed spectrum from the SAMI data cube.

B.3.2 Emission Line Fitting

Emission line maps of the most commonly used strong emission lines are provided in the value-added SAMI public release data. However, to ensure our emission line ratios are making self-consistent comparisons between strong-lines and the faint auroral lines, we perform new fits to these strong lines as well as the auroral lines. The emission line fitting procedure applied to each spaxel is as follows.

We first fit the Hα emission with a two-component Gaussian profile. We perform a χ^2-minimisation fit across the 6672 – 6698 Å wavelength range (rest-frame 6553 – 6578 Å at SAMI609396B catalog redshift). This range encompasses >99% of the Hα emission for the redshift range covered by the spaxels of this object while minimising contribution from nearby [N II] emission lines. The Hα emission is well modelled as a primary narrow component (median FWHM$_{H\alpha,\ nar}$ = 2.26 Å, median redshift z_{nar} = 0.018566) and a secondary broad component (median FWHM$_{H\alpha,\ brd}$ = 7.07 Å, median redshift z_{brd} = 0.018493) across the spatial extent of SAMI609396B.

We then fix the velocity and velocity dispersion for each of these two kinematic components to values obtained from Hα and simultaneously fit across the full optical wavelength range for broad- and narrow-component fluxes for each of the strong emission lines ([O II] $\lambda\lambda$3726,9, Hβ, [O III] $\lambda\lambda$4959,5007, [O I] λ6300, [N II] $\lambda\lambda$6548,83, Hα, [S II] λ6716, and [S II] λ6731).

The flux of each of these components is allowed to vary freely above a lower-bound of $f_{\rm comp} \geq 0$ with the exception of [O III]λ4959 and [N II]λ6548. The fluxes of each component of these lines are tied to the flux of the corresponding component of [O III] λ5007 and [N II] λ6583, respectively, according to following theoretical ratios: $f_{5007} = 3 \cdot f_{4959}$ and $f_{6583} = 2.9 \cdot f_{6548}$.

We calculate the uncertainty in the flux for each component by adding in quadrature the statistical error from the fit to an estimate of the uncertainty in the level of the continuum. This continuum uncertainty term is calculated as $\sigma_l = \sigma_c \cdot \sqrt{N + EW/\Delta}$ (Eq. 1 in Pérez-Montero 2017) where σ_c is the standard deviation in a 30 Å range near the emission line, selected to contain only continuum flux, N is the number of spectral pixels encompassed by the fit Gaussian, EW is the equivalent width of the line, and Δ is the spectral dispersion (Å/pixel). When considering the 'total' emission (i.e. sum

of both components), the adopted uncertainty is the uncertainty of both components summed in quadrature.

With the exception of [O I] λ6300, the faintest of these "strong-lines", the fits to these strong lines achieve summed component $S/N > 20$ across the entire spatial extent of interest and achieve $S/N > 50$ in over 95% of the spaxels. The [O I] λ6300 line fits achieve $S/N > 20$ in >95% of spaxels but only $S/N > 50$ in the brightest 8 spaxels. These two-component fits were visually inspected, adding confidence to the automatised algorithms.

The fainter emission lines do not present with sufficient signal-to-noise to be reliably modelled with two-component fits. Instead, to these fainter lines we make single component fits where the velocity and velocity dispersion are fixed to those derived for the dominant narrow component above. The faint lines for which single components are used include the three auroral emission lines visually identified to be present ([S II] $\lambda\lambda$4069,76, [O III] λ4363, and [S III] λ6312), and the fainter Balmer emission lines Hδ and Hγ. Additionally, we include the [Fe II] λ4360 emission line in our single component fit. The effect of the presence of this blended emission feature on the measured [O III] flux is discussed in detail in §B.3.3.

As before, these emission line fluxes are allowed to vary freely with constraint of $f \geq 0$ except that fluxes of the [S II] $\lambda\lambda$4069, 76 doublet are constrained such that $f_{4069} = 3 \cdot f_{4076}$. Uncertainties for each line are calculated in the same way as described for the individual components of the strong line fits described above. Once derived for each spaxel, these emission line fluxes are collated into 2D maps. Emission line fits for an example 1D spaxel spectrum are shown in Figure B.1.

As a final step, we correct these 2D emission line flux maps using values from the extinction correction map provided in the SAMI DR2 data, derived using spatially smoothed Balmer decrements ($f_{H\alpha}/f_{H\beta}$) to account for aliasing effects introduced by the SAMI observing process (refer to Green et al. 2018 and Medling et al. 2018 for details), assuming a Cardelli et al. (1989) extinction law with $R_V = 3.1$. Unless otherwise specified, the analysis of this paper is conducted using these reddening corrected emission line flux maps.

B.3.3 [FeII]λ4360 & [OIII]λ4363 blending

Several recent studies have identified an emission feature at λ4360 Å which may be blended with the [O III]λ4363 emission line, attributed to an [Fe II] emission line (Curti et al., 2017; Berg et al., 2020; Arellano-Córdova & Rodríguez, 2020). From visual inspection of 1D spectra, we identify that [O III]λ4363 emission often presents with an extended blue wing, which we attribute to blending with this [Fe II] λ4360 line.

We account for this by including an emission feature at this wavelength in our line-fitting routine whose flux is allowed to vary freely (Appendix B.3.2). As with all faint lines in the line-fitting, the velocity and velocity dispersion is tied to that of the narrow component identified for Hα. Visually, the fits obtained appear to model the emission features around λ4363 Å well. The median ratio between the [O III] and [Fe II] lines across the spatial extent with $S/N_{\lambda 4363} > 3$ is $f_{4363}/f_{4360} = 2.1$ with standard deviation $\sigma_{4363/4360} = 1.47$.

As a check of how reliable our [O III] flux measurements with this blended [Fe II] + [O III] profile are, we fit this wavelength region using three approaches and compare the results. The three approaches are:

1. *Standard fitting*: As described in Appendix B.3.2 where [Fe II] and [O III] components are simultaneously fit for.

2. *Naive single component*: A simple single component fit to [O III] across the wavelength range from 4345 – 4380 Å. No attempt is made to account for [Fe II] emission.

3. *Red wing single component*: A single component is fit to [O III], excluding pixels blueward of $\lambda = 4362.5 \times (1 + z_{\rm fit})$ Å, where $z_{\rm fit}$ is the redshift value obtained from the narrow-component fit to Hα for the spaxel in question. This should mask out spectral pixels with > 5 % contribution from [Fe II] emission.

In each approach, the velocity and velocity dispersion values are fixed to those obtained for the Hα narrow component, as in our standard fitting. The best-fit profiles from each of these approaches are shown in the top panel of Figure B.2. The bottom panel shows the distribution of values obtained for f_{alt}/f_{std} ratios, where f_{std} is the [O III] flux

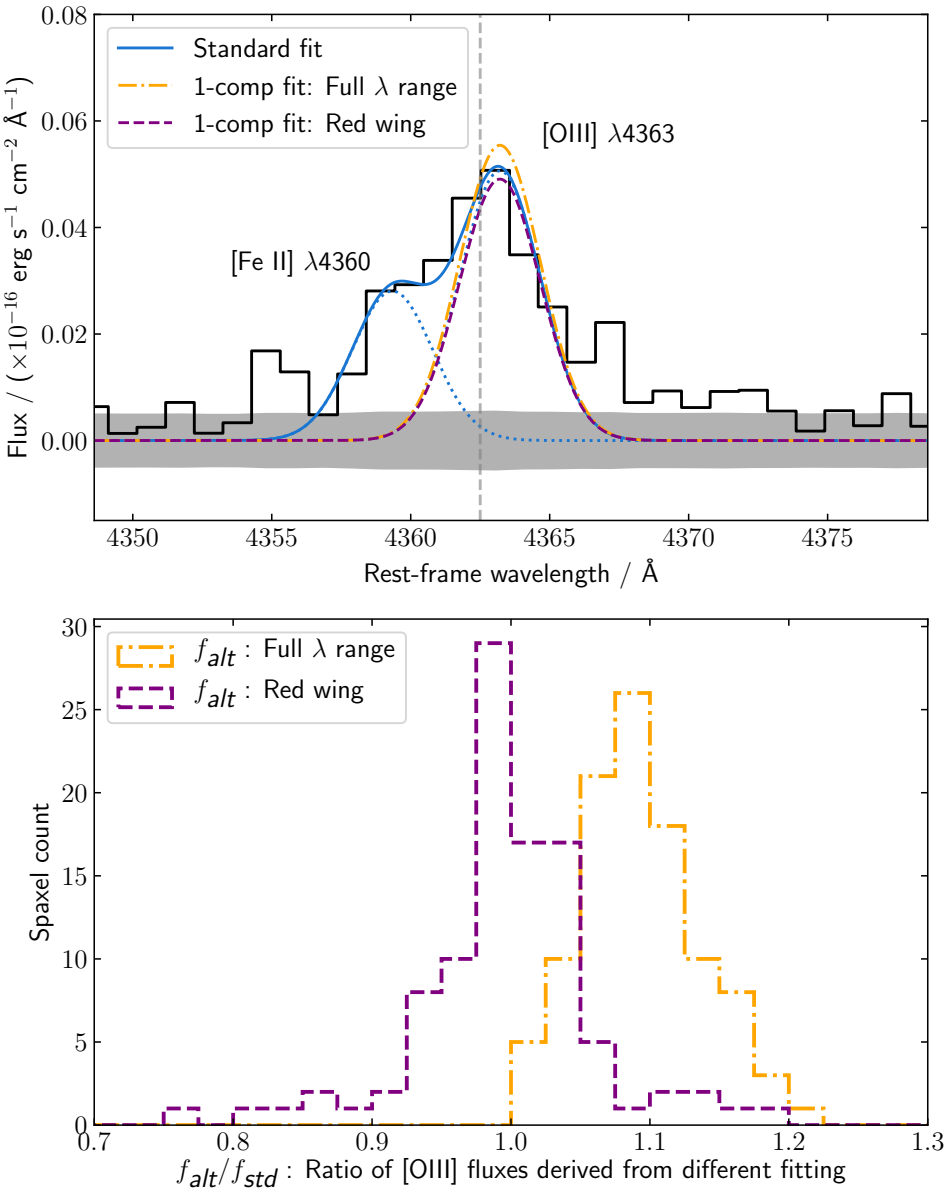

FIGURE B.2: *Top panel:* example of fits obtained to the region around rest-frame λ4363 when three different fitting methods are implemented. These methods are described in Appendix B.3.3. Method (i) is shown by blue solid line, with dotted lines showing each component. Method (ii) is shown by the orange dash-dotted line, while method (iii) is shown as the purple dashed line. The wavelength range used for method (iii) is that red-ward of 4362.5 Å (denoted by vertical dashed line). *Bottom panel:* Histogram of spaxels with $S/N_{\lambda4363} > 3$ of [OIII] flux measurements obtained from methods (ii) and (iii), shown relative to [OIII] flux obtained from method (i). As expected, when [FeII] emission is not accounted for as in (ii), [OIII] flux is systematically higher. While methods (iii) and (i) do not universally agree, the lack of systematic offset gives us confidence that either of these methods on average reliably account for [FeII] emission.

obtained from method (i) and f_{alt} is the flux from approaches (ii) and (iii). As expected, when no attempt is made to account for [Fe II] emission as in method (ii), the [O III] flux is systematically overestimated by around 10 % ($f_{alt}/f_{std} = 1.09 \pm 0.04$; orange dash-dotted line in bottom panel of Figure B.2). However, when applying method (iii), where the blue wing of [O III] is masked out, we see no significant systematic offset from our values obtained by method (i) ($f_{alt}/f_{std} = 0.99 \pm 0.07$; purple dashed line in bottom panel of Figure B.2). We note that there is large scatter in the distribution of f_{alt}/f_{std} for this latter case. This uncertainty of >10% is not unreasonable for observations with $S/N \sim 3 - 15$. We note also that when using these auroral line fluxes to derive electron temperature measurements, measurement uncertainties of this level are likely outweighed by modelling uncertainties (refer to § 4.3)

From this we conclude that for the [O III] /[Fe II] flux ratios we observe ($f_{4363}/f_{4360} \sim 2.1$), on average methods (i) and (iii) each suffer minimally from contamination by this blended emission feature at λ4360. [O III]λ4363 flux measurements quoted in other sections are those derived from the standard fitting routine (method (i)).

B.4 Implementing the method of Yates et al. (2020)

Measurements of electron temperature (T_e) are highly sought after in chemical abundance studies as they enable "direct-method" metallicity measurements to be made. Full application of the direct method to determine the total oxygen abundance requires measurements of both T_e([O III]) and T_e([O II]), such that both O^{2+}/H^+ and O^+/H^+ can be determined. However, given the faintness of the required auroral emission lines, often only T_e([O III]) can be measured directly. In this case, it is common to apply a "semi-direct" method. In this approach, the T_e([O II]) is indirectly determined from the measured T_e([O III]) via some assumed relation, allowing the O^+ abundance to be derived.

While it is common to assume a simple positive correlation between T_e([O II]) and T_e([O III]) (e.g. Izotov et al., 2006; López-Sánchez et al., 2012), Yates et al. (2020) (Y20 hereafter) find that this does a poor job of describing the observed scatter about this relation. Instead, Y20 highlight that at fixed metallicity, T_e([O II]) is anti-correlated with T_e([O III]), and that the general positive trend between T_e([O II]) and T_e([O III]) is

due to the fact that both correlate negatively with metallicity and in general will both be higher in lower metallicity systems.

Based on these observations, Yates et al. (2020) have outlined a new method for determining T_e([O II]) and metallicity in systems where only T_e([O III]) can be directly measured. Unlike previous semi-direct methods, in the Y20 method, T_e([O II]) and metallicity are solved for simultaneously. This differs from previous semi-direct methods in which T_e([O II]) is usually determined based on T_e([O III]) independently of metallicity. Metallicity is then subsequently determined using the value obtained for T_e([O II]).

We note that this method ("Y20 method") is separate to the empirical correction described in that same publication which we performed in § 4.4.1.2 ("Y20 correction"), and the Y20 empirical correction still needs to be applied to the metallicities arising from the Y20 method.

In Section 4.4 we applied a simple semi-direct method to our SAMI609396B data in which T_e([O II]) was determined from T_e([O III]) using the calibration of López-Sánchez et al. (2012) (Eq. 4.6 in § 4.4.1) and then the metallicity subsequently determined accordingly (that method was referred to as $Z_{\mathrm{Te;\,LS12}}$ throughout this work). Here, we additionally apply the Y20 method to our SAMI609396B data. We find the results of the Y20 method to be double-valued (Figure B.3). The "upper branch" largely agrees with our $Z_{\mathrm{Te;\,LS12}}$ metallicities within ∼0.15 dex (median absolute offset is 0.06 dex), however the "lower branch" gives starkly different values. Spatial metallicity trends arising from each branch of this method differ noticeably (Figure B.4), making it difficult to draw conclusions without further characterisation of the behaviour of each branch in a larger data set. We describe the details of our implementation of the Y20 method below.

B.4.1 Basis of the Yates Method

The Y20 method differs from other semi-direct methods in that T_e([O II]) and metallicity (Z_{Te}) are evaluated simultaneously in order to account for the interdependence of these parameters at fixed T_e([O III]); whereas typically semi-direct methods have involved inferring a T_e([O II]) value from a T_e([O III]) measurement via a fixed relation and then subsequently determining the metallicity using this value.

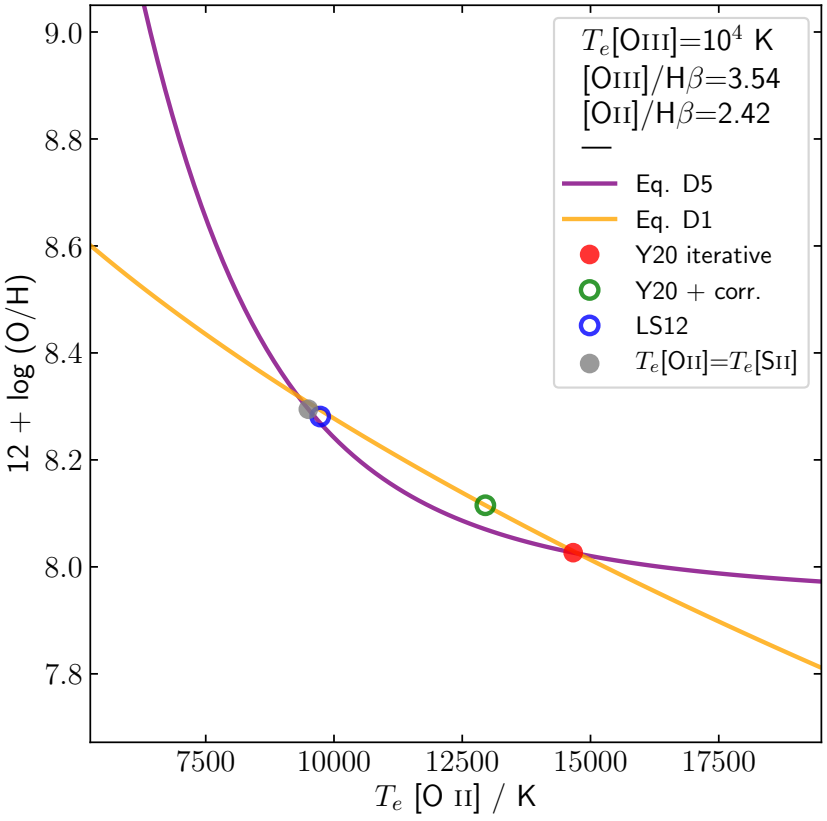

FIGURE B.3: Solving for the intersection of equations B.1 and B.5 from the Y20 method yields two solutions: an 'upper-branch' solution on the left-hand side with a high metallicity, and a 'lower-branch' value on the right-hand side with a low metallicity. Here, the relationship between T_e([O II]) and $Z_{\rm Te}$ is shown for fixed values of T_e([O III]), [O III]/Hβ, and [O II]/Hβ measured from a typical spaxel in SAMI609396B (see legend). The purple line shows oxygen abundance determined according to equations B.3 – B.5 for various values of T_e([O II]), while the orange line shows Yates et al. (2020) metallicity-dependent fit to the T_e([O II]) – T_e([O III]) relation at a fixed value of T_e([O III]). The Y20 iterative method as originally outlined favours the lower branch solution (red circle). The Y20 correction increases the metallicity (green open circle), although not to the value of the upper branch. Both $Z_{\rm Te;\,LS12}$ and $Z_{\rm Te;\,SII}$ (refer to § 4.4; T_e([S II]) = 9500 K) fall quite close to the upper branch solution (open blue circle and filled grey circle respectively).

The Y20 method centres on a metallicity-dependent fit to the $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation, outlined as follows:

$$T_e([\mathrm{O\,II}]) = \frac{a(Z_{\mathrm{Te}})^2}{2 \cdot T_e([\mathrm{O\,III}])}, \tag{B.1}$$

where

$$a = -12030.22 \cdot Z_{\mathrm{Te}} + 113720.75. \tag{B.2}$$

This can be solved simultaneously with the following equations which determine oxygen abundance from measured $[\mathrm{O\,II}]/\mathrm{H}\beta$ and $[\mathrm{O\,III}]/\mathrm{H}\beta$ line ratios given values for $T_e([\mathrm{O\,II}])$ and $T_e([\mathrm{O\,III}])$:

$$O^+/H^+ = \frac{[\mathrm{O\,II}]\lambda\lambda 3726,9}{\mathrm{H}\beta}\ g_1\ \alpha_{\mathrm{H}\beta}\ \sqrt{T_e[\mathrm{O\,II}]} \times \exp[E_{12}/kT_e([\mathrm{O\,II}])] \times \frac{\beta}{E_{12}\Upsilon_{12}} \tag{B.3}$$

$$O^{++}/H^+ = \frac{[\mathrm{O\,III}]\lambda\lambda 4959,5007}{\mathrm{H}\beta}\ g_1\ \alpha_{\mathrm{H}\beta}\ \sqrt{T_e[\mathrm{O\,III}]} \times \exp[E_{12}/kT_e([\mathrm{O\,III}])] \times \frac{\beta}{E_{12}\Upsilon_{12}} \tag{B.4}$$

$$Z_{\mathrm{Te}} \equiv 12 + \log(O^+/H^+ + O^{++}/H^+). \tag{B.5}$$

The reader is referred to Yates et al. (2020) and Nicholls et al. (2014) for more details on these equations including the values and calculations of various parameters used.

B.4.2 Two-valued solution of the Yates method

In Yates et al. (2020), the authors propose solving these equations with fixed point iteration, however we instead propose numerically solving for the intersection of equations B.1 & B.5. This preference is based on our observation that the relations described in Equations B.1 & B.5 in fact yield two solutions within the range of what could be considered physically reasonable.[1]

[1] More generally, it is possible that for some observations there will be no solution. In these cases, solving via the iterative method may be preferable (Yates 2020; private communication)

This is illustrated in Figure B.3 for an example typical SAMI609396B spaxel with $T_e([\mathrm{O\,III}]) = 10^4$ K, and [O III]/Hβ and [O II]/Hβ line ratios of 3.54 and 2.42 respectively. For this example spaxel it can be seen that two possible solutions exist: (1) at $T_e([\mathrm{O\,II}]) = 9{,}342$ K and $Z_{\mathrm{Te}} = 8.31$, and (2) at $T_e([\mathrm{O\,II}]) = 14{,}662$ K and $Z_{\mathrm{Te}} = 8.03$. Neither of these solutions is physically implausible and while solution (1) would fall in a more densely populated region of the $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation as shown in Yates et al. (2020) (refer to Figure 5 in that paper), observed points comparable to solution (2) are found in their sample too.

Indeed, beyond the single example shown in Figure B.3, we find that all spaxels in SAMI609396B with at least one solution have precisely two. How then should we decide which of these two solutions to adopt?

The blue open circle in Figure B.3 shows that the $T_e([\mathrm{O\,II}])$ obtained via the López-Sánchez et al. (2012) $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relation (refer to Eq. 4.6 in § 4.4.1) agrees quite well with the lower valued $T_e([\mathrm{O\,II}])$ solution. However, this disregards the point of the Y20 method and findings presented in this work: that assuming a simple fixed relationship between $T_e([\mathrm{O\,II}])$ and $T_e([\mathrm{O\,III}])$ can be misleading.

The Y20 iterative method as originally applied in Yates et al. (2020) selects for the "lower branch" (red circle in Fig B.3). However, subsequent application of the Y20 correction serves to shift this point to a slightly higher metallicity, partly toward the upper branch solution (green open circle in Fig B.3).

To investigate this further, the original Y20 sample was revisited with the two-valued solution in mind (Yates 2020; private communication). It was found that targets where the direct[2] metallicity was closer to the "upper branch" solution were often targets with an $O32$ value below the threshold value for which the Y20 correction should be applied ($O32 \leq 0.29$). Thus, this $O32$ threshold could be used to distinguish between the lower and upper branches. All spaxels across the spatial extent of SAMI609396B fall in this category with $O32 \leq 0.29$ (refer to Fig 4.8), meaning we would adopt the upper branch value under this scheme, rather than the lower branch value favoured by the original iterative implementation.

[2]Here "direct" metallicity refers to a metallicity in which both $T_e([\mathrm{O\,II}])$ and $T_e([\mathrm{O\,III}])$ have been directly measured with auroral lines.

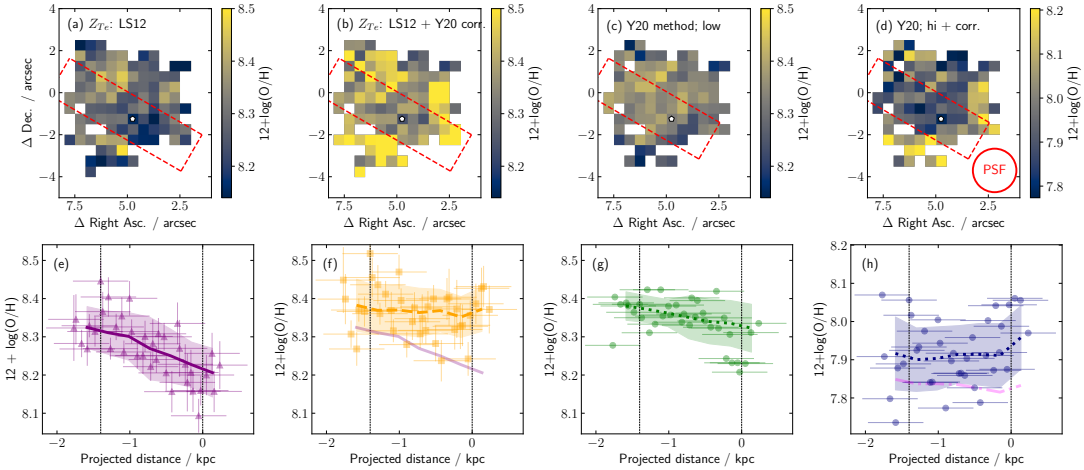

FIGURE B.4: Comparison between the $Z_{\mathrm{Te;\,LS12}}$ method described in § 4.4.1 and the two solutions arising from the Y20 method. Far left and centre-left columns are simply reproduced from Figure 4.5. The centre-right column shows the Y20 upper branch metallicities, while the far-right column shows corrected Y20 lower branch metallicities (the uncorrected lower branch metallicity trend is shown in panel h as the magenta dot-dashed line). Note: We do not take the step of propagating measurement errors through the Y20 method. However, the input measurement uncertainties are from the same source as those in panels (e) and (f) and thus would likely be comparable.

While we do not directly measure T_e([O II]), in § 4.3.3 we derived T_e([S II]) from the [S II] $\lambda\lambda$4069, 76 / $\lambda\lambda$6716, 31 line ratio (refer to Figure 4.4). Panel (e) of Figure 4.4 shows that T_e([S II]) < T_e([O III]) across the majority of spaxels in SAMI609396B, with a median value of T_e([S II]) = 9,295 K. Given that previous studies have found that T_e([S II]) and T_e([O II]) are often in general agreement (e.g. Croxall et al., 2016), we consider that this additionally supports our selection of the upper-branch value. We note, however, that large scatter is known to exist in both the T_e([S II]) – T_e([O II]) and T_e([S II]) – T_e([O III]) relations.

B.4.3 Comparison between Yates method and LS12 method

In line with our methods outlined in Appendix B.4.2 above, we consider two versions of the Y20 method: one in which we adopt the "upper branch" solution to the Y20 method, and another where we adopt the "lower branch" and apply the Y20 empirical correction. We find that the resulting T_e([O II]) and Z_{Te} maps are smooth with no unexpectedly large variations observed between pairs of adjacent spaxels.

In Figure B.4 we compare the metallicities obtained from these Y20 methods with those obtained in § 4.4.1 from the $Z_{\mathrm{Te;\,LS12}}$ method. Metallicity maps and spatial trends in

the far- and centre-left columns of Figure B.4 are simply reproduced from Figure 4.5 and show the $Z_{\mathrm{Te;\ LS12}}$ method with and without the Y20 empirical correction. The centre-right column shows the Y20 upper branch metallicities, while the far-right column shows corrected Y20 lower branch metallicities (additionally, the uncorrected lower branch metallicity trend is shown in panel h as the magenta dot-dashed line).

We first note that, even after the Y20 correction, the lower branch metallicities are significantly lower than any of our other semi-direct methods.[3] The large scatter in the values obtained makes it difficult to determine the spatial metallicity trend from this method, however, qualitatively it does seem to be broadly consistent with the trend seen in the $Z_{\mathrm{Te;\ Y20}}$ method applied in § 4.4.1.2.

The normalisation of the upper branch values is in much better agreement with other semi-direct methods. The effect on the spatial trend is less clear: it appears somewhat flattened compared to $Z_{\mathrm{Te;\ LS12}}$, however if a gradient were to be computed it would likely depend strongly on a cluster of lower metallicity points with projected distance $r \approx 0$ kpc.

Yates et al. (2020) showed that the semi-direct abundance deficit at low O^{++}/O^{+} (which the Y20 correction aims to address) is present across all $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relations considered in that work, including LS12 (Figure 6 in Yates et al. 2020). In particular, they show in detail its effect on the Y20 lower branch metallicities. Although it seems that low O^{++}/O^{+} values seem to correlate with an increased preference for the upper branch metallicity solution, it is currently unclear whether variations in O^{++}/O^{+} result in semi-direct metallicity biases in a similar way to that observed by Yates et al. (2020) with respect to the lower branch solution and other $T_e([\mathrm{O\,II}])$ – $T_e([\mathrm{O\,III}])$ relations. In the context of SAMI609396B where we have shown large scale variations in the $[\mathrm{O\,III}]/[\mathrm{O\,II}]$ ratio, the existence of such a bias could affect our interpretation of the spatial metallicity trend resulting from this upper branch solution. Addressing this issue would require a more detailed analysis of a more extensive sample (e.g. the Yates et al. (2020) sample) and is beyond the scope of this paper.

In summary, we find that the Y20 semi-direct method as originally proposed favours a similar spatial gradient to that of $Z_{\mathrm{Te;\ Y20}}$ after the application of the Y20 correction,

[3] And, indeed, strong-line methods; although some degree of offset is expected there (e.g. Kewley & Ellison, 2008)

albeit at a much lower normalisation. After identifying the two-valued nature of these relations, we found that adopting the upper branch values resulted in normalisation that agreed much better with other methods. We defer commenting on the spatial trend arising from these upper branch values to a later study, after the two-valued nature of the Y20 relations has been examined in more detail.

Overall, these findings do not alter the main conclusion of this work: that assumptions around the temperature structure of H II regions can have a significant impact on measured spatial metallicity trends in IFU observations of galaxies.

Bibliography

Abbott B. P., et al., 2017, PhRvL, 119, 161101

Adelberger K. L., Steidel C. C., Giavalisco M., Dickinson M., Pettini M., Kellogg M., 1998, ApJ, 505, 18

Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005, ApJ, 619, 697

Allen M. G., Groves B. A., Dopita M. A., Sutherland R. S., Kewley L. J., 2008, ApJS, 178, 20

Allen J. T., et al., 2015, MNRAS, 446, 1567

Amorín R., et al., 2014, A&A, 568, L8

Andreani P., Cristiani S., Lucchin F., Matarrese S., Moscardini L., 1994, ApJ, 430, 458

Andrews B. H., Martini P., 2013, ApJ, 765, 140

Arellano-Córdova K. Z., Rodríguez M., 2020, arXiv e-prints, p. arXiv:2006.07393

Baldry I. K., et al., 2012, MNRAS, 421, 621

Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5

Barbary K., 2016, extinction v0.3.0, doi:10.5281/zenodo.804967, `https://doi.org/10.5281/zenodo.804967`

Barone-Nugent R. L., et al., 2014, ApJ, 793, 17

Bekki K., 2009, MNRAS, 399, 2221

Belfiore F., et al., 2017, MNRAS, 469, 151

Belli S., Jones T., Ellis R. S., Richard J., 2013, ApJ, 772, 141

Benítez N., 2000, ApJ, 536, 571

Benítez N., et al., 2004, ApJS, 150, 1

Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38

Berg D. A., et al., 2012, ApJ, 754, 98

Berg D. A., Skillman E. D., Garnett D. R., Croxall K. V., Marble A. R., Smith J. D., Gordon K., Kennicutt Robert C. J., 2013, ApJ, 775, 128

Berg D. A., Skillman E. D., Croxall K. V., Pogge R. W., Moustakas J., Johnson-Groh M., 2015, ApJ, 806, 16

Berg D. A., Pogge R. W., Skillman E. D., Croxall K. V., Moustakas J., Rogers N. S. J., Sun J., 2020, arXiv e-prints, p. arXiv:2001.05002

Bertin E., Arnouts S., 1996a, A&AS, 117, 393

Bertin E., Arnouts S., 1996b, A&AS, 117, 393

Bian F., Kewley L. J., Dopita M. A., Juneau S., 2016, ApJ, 822, 62

Bian F., Kewley L. J., Dopita M. A., 2018, ApJ, 859, 175

Bian F., Kewley L. J., Groves B., Dopita M. A., 2020, MNRAS, 493, 580

Blanc G. A., et al., 2008, ApJ, 681, 1099

Blanc G. A., Kewley L., Vogt F. P. A., Dopita M. A., 2015, ApJ, 798, 99

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Borthakur S., et al., 2016, ApJ, 833, 259

Bothwell M. S., Maiolino R., Kennicutt R., Cresci G., Mannucci F., Marconi A., Cicone C., 2013, MNRAS, 433, 1425

Bradley L. D., et al., 2012, ApJ, 760, 108

Bradley L., et al., 2019, astropy/photutils: v0.7.2, doi:10.5281/zenodo.3568287, `https://doi.org/10.5281/zenodo.3568287`

Bresolin F., 2019, MNRAS, 488, 3826

Bresolin F., Kennicutt R. C., 2015, MNRAS, 454, 3664

Bresolin F., Garnett D. R., Kennicutt Robert C. J., 2004, ApJ, 615, 228

Bresolin F., Ryan-Weber E., Kennicutt R. C., Goddard Q., 2009, ApJ, 695, 580

Brooks A. M., Governato F., Booth C. M., Willman B., Gardner J. P., Wadsley J., Stinson G., Quinn T., 2007, ApJL, 655, L17

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bryant J. J., et al., 2015, MNRAS, 447, 2857

Budavári T., et al., 2003, ApJ, 595, 59

Bundy K., et al., 2015, ApJ, 798, 7

Calabrò A., et al., 2017, A&A, 601, A95

Calura F., Pipino A., Chiappini C., Matteucci F., Maiolino R., 2009, A&A, 504, 373

Calvi V., et al., 2016, ApJ, 817, 120

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Cameron E., Carollo C. M., Oesch P. A., Bouwens R. J., Illingworth G. D., Trenti M., Labbé I., Magee D., 2011, ApJ, 743, 146

Campbell A., Terlevich R., Melnick J., 1986, MNRAS, 223, 811

Cappellari M., 2017, MNRAS, 466, 798

Cappellari M., Verolme E. K., van der Marel R. P., Verdoes Kleijn G. A., Illingworth G. D., Franx M., Carollo C. M., de Zeeuw P. T., 2002, ApJ, 578, 787

Cardamone C., et al., 2009, MNRAS, 399, 1191

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Carton D., et al., 2018, MNRAS, 478, 4293

Casertano S., et al., 2000, AJ, 120, 2747

Chabrier G., 2003, PASP, 115, 763

Chisholm J., Tremonti C., Leitherer C., 2018, MNRAS, 481, 1690

Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926

Coil A. L., Mendez A. J., Eisenstein D. J., Moustakas J., 2017, ApJ, 838, 87

Colless M., et al., 2001, MNRAS, 328, 1039

Conselice C. J., Wilkinson A., Duncan K., Mortlock A., 2016, ApJ, 830, 83

Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839

Croom S. M., et al., 2012, MNRAS, 421, 872

Croton D. J., et al., 2006, MNRAS, 365, 11

Croxall K. V., Pogge R. W., Berg D. A., Skillman E. D., Moustakas J., 2015, ApJ, 808, 42

Croxall K. V., Pogge R. W., Berg D. A., Skillman E. D., Moustakas J., 2016, ApJ, 830, 4

Curti M., Cresci G., Mannucci F., Marconi A., Maiolino R., Esposito S., 2017, MNRAS, 465, 1384

Curti M., Mannucci F., Cresci G., Maiolino R., 2020a, MNRAS, 491, 944

Curti M., et al., 2020b, MNRAS, 492, 821

Curtis H. D., 1917, PASP, 29, 206

Curtis H. D., 1918, Publications of Lick Observatory, 13, 9

Cyburt R. H., Fields B. D., Olive K. A., Yeh T.-H., 2016, Reviews of Modern Physics, 88, 015004

D'Agostino J. J., Poetrodjojo H., Ho I. T., Groves B., Kewley L., Madore B. F., Rich J., Seibert M., 2018, MNRAS, 479, 4907

Daddi E., Cimatti A., Renzini A., Fontana A., Mignoli M., Pozzetti L., Tozzi P., Zamorani G., 2004, ApJ, 617, 746

Dalcanton J. J., Yoachim P., Bernstein R. A., 2004, ApJ, 608, 189

Davé R., Hellsten U., Hernquist L., Katz N., Weinberg D. H., 1998, ApJ, 509, 661

Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98

Davies R. L., Rich J. A., Kewley L. J., Dopita M. A., 2014a, MNRAS, 439, 3835

Davies R. L., Kewley L. J., Ho I. T., Dopita M. A., 2014b, MNRAS, 444, 3961

Davies R. L., et al., 2016, MNRAS, 462, 1616

Davies R. L., et al., 2017, MNRAS, 470, 4974

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Deharveng L., Peña M., Caplan J., Costero R., 2000, MNRAS, 311, 329

Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, ApJ, 142, 414

Dodelson S., 2004, Phys. Rev. D, 70, 023009

Dopita M. A., Sutherland R. S., 2003, Astrophysics of the diffuse universe

Dopita M. A., Sutherland R. S., 2017, ApJS, 229, 35

Dopita M. A., Kewley L. J., Heisler C. A., Sutherland R. S., 2000, ApJ, 542, 224

Dopita M. A., Sutherland R. S., Nicholls D. C., Kewley L. J., Vogt F. P. A., 2013, ApJS, 208, 10

Dopita M. A., Kewley L. J., Sutherland R. S., Nicholls D. C., 2016, Ap&SS, 361, 61

Edmunds M. G., Greenhow R. M., 1995, MNRAS, 272, 241

Efstathiou G., 1995, MNRAS, 276, 1425

Efstathiou G., Kaiser N., Saunders W., Lawrence A., Rowan-Robinson M., Ellis R. S., Frenk C. S., 1990, MNRAS, 247, 10P

Einstein A., 1916, Annalen der Physik, 354, 769

Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, PASA, 34, e058

Ellison S. L., Patton D. R., Simard L., McConnachie A. W., 2008, ApJL, 672, L107

Erb D. K., 2008, ApJ, 674, 151

Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006a, ApJ, 644, 813

Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006b, ApJ, 646, 107

Ferland G. J., et al., 2017, Rev. Mexicana Astron. Astrofis., 53, 385

Finlator K., 2017, Gas Accretion and Galactic Chemical Evolution: Theory and Observations. Springer, p. 221, doi:10.1007/978-3-319-52512-9˙10

Fraternali F., Binney J. J., 2008, MNRAS, 386, 935

García-Díaz M. T., Henney W. J., López J. A., Doi T., 2008, Rev. Mexicana Astron. Astrofis., 44, 181

Garstang R. H., 1951, MNRAS, 111, 115

Gibson B. K., Pilkington K., Brook C. B., Stinson G. S., Bailin J., 2013, A&A, 554, A47

Glover S., 2005, SSRv, 117, 445

Green A. W., et al., 2014, MNRAS, 437, 1070

Green A. W., et al., 2018, MNRAS, 475, 716

Groves B., Brinchmann J., Walcher C. J., 2012, MNRAS, 419, 1402

Gunn J. E., Gott J. Richard I., 1972, ApJ, 176, 1

Harikane Y., et al., 2018, PASJ, 70, S11

Harris C. R., et al., 2020, Nature, 585, 357

Hayashi M., Shimasaku K., Motohara K., Yoshida M., Okamura S., Kashikawa N., 2007, ApJ, 660, 72

Heckman T. M., 2002, in Mulchaey J. S., Stocke J. T., eds, Astronomical Society of the Pacific Conference Series Vol. 254, Extragalactic Gas at Low Redshift. p. 292 (arXiv:astro-ph/0107438)

Heckman T. M., Armus L., Miley G. K., 1990, ApJS, 74, 833

Heckman T. M., Lehnert M. D., Strickland D. K., Armus L., 2000, ApJS, 129, 493

Heckman T. M., et al., 2005, ApJL, 619, L35

Heger A., Woosley S. E., 2002, ApJ, 567, 532

Hemler Z. S., et al., 2020, arXiv e-prints, p. arXiv:2007.10993

Henry A., Martin C. L., Finlator K., Dressler A., 2013a, ApJ, 769, 148

Henry A., et al., 2013b, ApJL, 776, L27

Hermit S., Santiago B. X., Lahav O., Strauss M. A., Davis M., Dressler A., Huchra J. P., 1996, MNRAS, 283, 709

Herwig F., 2005, ARA&A, 43, 435

Hillebrandt W., Niemeyer J. C., 2000, ARA&A, 38, 191

Hirtenstein J., et al., 2019, ApJ, 880, 54

Ho I. T., et al., 2014, MNRAS, 444, 3894

Ho I. T., Kudritzki R.-P., Kewley L. J., Zahid H. J., Dopita M. A., Bresolin F., Rupke D. S. N., 2015, MNRAS, 448, 2030

Ho I. T., et al., 2016, MNRAS, 457, 1257

Ho I. T., et al., 2019, ApJL, 885, L31

Hoag A., et al., 2019, ApJ, 878, 12

Hobbs A., Read J., Nicola A., 2015, MNRAS, 452, 3593

Hollenbach D., McKee C. F., 1979, ApJS, 41, 555

Hopkins P. F., Quataert E., Murray N., 2012, MNRAS, 421, 3522

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hunter J. D., 2007, Computing in Science & Engineering, 9, 90

Ilbert O., et al., 2006, A&A, 457, 841

Ishikawa S., Kashikawa N., Toshikawa J., Tanaka M., Hamana T., Niino Y., Ichikawa K., Uchiyama H., 2017, ApJ, 841, 8

Izotov Y. I., Stasińska G., Meynet G., Guseva N. G., Thuan T. X., 2006, A&A, 448, 955

Janka H.-T., 2012, Annual Review of Nuclear and Particle Science, 62, 407

Jones T., Ellis R., Jullo E., Richard J., 2010, ApJL, 725, L176

Jones T., Ellis R. S., Richard J., Jullo E., 2013, ApJ, 765, 48

Jones T., et al., 2015a, AJ, 149, 107

Jones T., Martin C., Cooper M. C., 2015b, ApJ, 813, 126

Jones A., et al., 2017, A&A, 599, A141

Jones T., Sanders R., Roberts-Borsani G., Ellis R. S., Laporte N., Treu T., Harikane Y., 2020, arXiv e-prints, p. arXiv:2006.02447

Kacprzak G. G., Churchill C. W., Nielsen N. M., 2012, ApJL, 760, L7

Kaiser N., 1984, ApJL, 284, L9

Karakas A. I., 2010, MNRAS, 403, 1413

Karlsson T., Bromm V., Bland-Hawthorn J., 2013, Reviews of Modern Physics, 85, 809

Kauffmann G., et al., 2003, MNRAS, 346, 1055

Kent S. M., Gunn J. E., 1982, AJ, 87, 945

Kewley L. J., Dopita M. A., 2002, ApJS, 142, 35

Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183

Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Kewley L. J., Rupke D., Zahid H. J., Geller M. J., Barton E. J., 2010, ApJL, 721, L48

Kewley L. J., Maier C., Yabe K., Ohta K., Akiyama M., Dopita M. A., Yuan T., 2013, ApJL, 774, L10

Kewley L. J., Nicholls D. C., Sutherland R. S., 2019a, ARA&A, 57, 511

Kewley L. J., Nicholls D. C., Sutherland R., Rigby J. R., Acharya A., Dopita M. A., Bayliss M. B., 2019b, ApJ, 880, 16

Kobayashi C., Karakas A. I., Lugaro M., 2020, ApJ, 900, 179

Kobulnicky H. A., Kewley L. J., 2004, ApJ, 617, 240

Koekemoer A. M., Fruchter A. S., Hook R. N., Hack W., 2003, in Arribas S., Koekemoer A., Whitmore B., eds, HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System. p. 337

Köppen J., Weidner C., Kroupa P., 2007, MNRAS, 375, 673

Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221

Lara-López M. A., et al., 2010, A&A, 521, L53

Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692

Lee K.-S., Giavalisco M., Gnedin O. Y., Somerville R. S., Ferguson H. C., Dickinson M., Ouchi M., 2006, ApJ, 642, 63

Leethochawalit N., Jones T. A., Ellis R. S., Stark D. P., Richard J., Zitrin A., Auger M., 2016, ApJ, 820, 84

Leitherer C., Robert C., Drissen L., 1992, ApJ, 401, 596

Leitherer C., Ekström S., Meynet G., Schaerer D., Agienko K. B., Levesque E. M., 2014, ApJS, 212, 14

Lennon D. J., Burke V. M., 1994, A&AS, 103, 273

Lequeux J., Peimbert M., Rayo J. F., Serrano A., Torres-Peimbert S., 1979, A&A, 80, 155

Leung G. C. K., Coil A. L., Rupke D. S. N., Perrotta S., 2020, arXiv e-prints, p. arXiv:2011.09587

Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21

Li Y., Bresolin F., Kennicutt Robert C. J., 2013, ApJ, 766, 17

Licquia T. C., Newman J. A., 2015, ApJ, 806, 96

Lin L., et al., 2012, ApJ, 756, 71

Linde A. D., 1983, Physics Letters B, 129, 177

Lintott C., et al., 2011, MNRAS, 410, 166

López-Sánchez Á. R., Dopita M. A., Kewley L. J., Zahid H. J., Nicholls D. C., Scharwächter J., 2012, MNRAS, 426, 2630

López-Sanjuan C., et al., 2014, A&A, 564, A127

López-Sanjuan C., et al., 2015, A&A, 582, A16

Luridiana V., Morisset C., Shaw R. A., 2015, A&A, 573, A42

Ly C., Malhotra S., Malkan M. A., Rigby J. R., Kashikawa N., de los Reyes M. A., Rhoads J. E., 2016a, ApJS, 226, 5

Ly C., Malkan M. A., Rigby J. R., Nagao T., 2016b, ApJ, 828, 67

Ma X., Hopkins P. F., Feldmann R., Torrey P., Faucher-Giguère C.-A., Kereš D., 2017, MNRAS, 466, 4780

Mac Low M. M., Shull J. M., 1986, ApJ, 302, 585

Machacek M. E., Bryan G. L., Abel T., 2001, ApJ, 548, 509

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Madgwick D. S., et al., 2003, MNRAS, 344, 847

Magrini L., Coccato L., Stanghellini L., Casasola V., Galli D., 2016, A&A, 588, A91

Maiolino R., Mannucci F., 2019, A&A Rev., 27, 3

Maiolino R., et al., 2008, A&A, 488, 463

Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115

Marasco A., Fraternali F., Binney J. J., 2012, MNRAS, 419, 1107

Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, ApJ, 701, 1765

Marino R. A., et al., 2013, A&A, 559, A114

Maseda M. V., et al., 2014, ApJ, 791, 17

Mason C. A., Trenti M., Treu T., 2015, ApJ, 813, 21

Mason C. A., et al., 2017, ApJ, 838, 14

Masters D., et al., 2014, ApJ, 785, 153

Mather J. C., et al., 1990, ApJL, 354, L37

Matteucci F., Recchi S., 2001, ApJ, 558, 351

McCall M. L., Rybski P. M., Shields G. A., 1985, ApJS, 57, 1

McCracken H. J., et al., 2010, ApJ, 708, 202

McGaugh S. S., 1991, ApJ, 380, 140

Medling A. M., et al., 2018, MNRAS, 475, 5194

Meneux B., et al., 2009, A&A, 505, 463

Millington S. J. C., Peach J. V., 1986, MNRAS, 221, 15

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Mo H. J., White S. D. M., 2002, MNRAS, 336, 112

Morales-Luis A. B., Pérez-Montero E., Sánchez Almeida J., Muñoz-Tuñón C., 2014, ApJ, 797, 81

Mostek N., Coil A. L., Cooper M., Davis M., Newman J. A., Weiner B. J., 2013, ApJ, 767, 89

Moustakas J., Kennicutt Robert C. J., Tremonti C. A., Dale D. A., Smith J.-D. T., Calzetti D., 2010, ApJS, 190, 233

Naoz S., Noter S., Barkana R., 2006, MNRAS, 373, L98

Newville M., Stensitzki T., Allen D. B., Ingargiola A., 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, doi:10.5281/zenodo.11813, https://doi.org/10.5281/zenodo.11813

Nicholls D. C., Dopita M. A., Sutherland R. S., Jerjen H., Kewley L. J., Basurah H., 2014, ApJ, 786, 155

Nicholls D. C., Kewley L. J., Sutherland R. S., 2020, PASP, 132, 033001

Nishimura N., Takiwaki T., Thielemann F.-K., 2015, ApJ, 810, 109

Norberg P., et al., 2001, MNRAS, 328, 64

Norberg P., et al., 2002, MNRAS, 332, 827

Oesch P. A., et al., 2016, ApJ, 819, 129

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713

Olive K. A., Steigman G., Walker T. P., 2000, Physics Reports, 333, 389

Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312

Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei

Ouchi M., et al., 2004, ApJ, 611, 685

Ouchi M., et al., 2005, ApJL, 635, L117

Palay E., Nahar S. N., Pradhan A. K., Eissner W., 2012, MNRAS, 423, L35

Patrício V., Christensen L., Rhodin H., Cañameras R., Lara-López M. A., 2018, MNRAS, 481, 3520

Pearson W. J., et al., 2019, A&A, 631, A51

Peebles P. J. E., 1980, The large-scale structure of the universe

Peimbert M., 1967, ApJ, 150, 825

Peimbert M., Peimbert A., Delgado-Inglada G., 2017, PASP, 129, 082001

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Pérez-Montero E., 2014, MNRAS, 441, 2663

Pérez-Montero E., 2017, PASP, 129, 043001

Pérez-Montero E., Contini T., 2009, MNRAS, 398, 949

Perlmutter S., et al., 1999, ApJ, 517, 565

Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59

Pettini M., Madau P., Bolte M., Prochaska J. X., Ellison S. L., Fan X., 2003, ApJ, 594, 695

Pillepich A., Madau P., Mayer L., 2015, ApJ, 799, 184

Pilyugin L. S., Thuan T. X., 2005, ApJ, 631, 231

Pilyugin L. S., Grebel E. K., Mattsson L., 2012, MNRAS, 424, 2316

Planck Collaboration et al., 2016, A&A, 594, A1

Poetrodjojo H., et al., 2018, MNRAS, 479, 5235

Poetrodjojo H., D'Agostino J. J., Groves B., Kewley L., Ho I. T., Rich J., Madore B. F., Seibert M., 2019, MNRAS, 487, 79

Postman M., et al., 2012, ApJS, 199, 25

Proxauf B., Öttl S., Kimeswenger S., 2014a, A&A, 561, A10

Proxauf B., Öttl S., Kimeswenger S., 2014b, A&A, 561, A10

Qiu Y., et al., 2018, MNRAS, 481, 4885

Renzini A., Peng Y.-j., 2015, ApJL, 801, L29

Rich J. A., Dopita M. A., Kewley L. J., Rupke D. S. N., 2010, ApJ, 721, 505

Richard J., Jones T., Ellis R., Stark D. P., Livermore R., Swinbank M., 2011, MNRAS, 413, 643

Riess A. G., et al., 1998, AJ, 116, 1009

Robertson B. E., 2010, ApJL, 716, L229

Rood H. J., 1979, ApJ, 232, 699

Rupke D., 2018, Galaxies, 6, 138

Salpeter E. E., 1955, ApJ, 121, 161

Sánchez Almeida J., Elmegreen B. G., Muñoz-Tuñón C., Elmegreen D. M., 2014, A&A Rev., 22, 71

Sánchez S. F., et al., 2012, A&A, 538, A8

Sánchez S. F., et al., 2019, MNRAS, 484, 3042

Sanders R. L., et al., 2015, ApJ, 799, 138

Sanders R. L., et al., 2016, ApJ, 816, 23

Sanders R. L., Shapley A. E., Zhang K., Yan R., 2017, ApJ, 850, 136

Sanders R. L., et al., 2020a, arXiv e-prints, p. arXiv:2009.07292

Sanders R. L., et al., 2020b, MNRAS, 491, 1427

Sato T., Sawicki M., Arcila-Osejo L., 2014, MNRAS, 443, 2661

Savaglio S., et al., 2005, ApJ, 635, 260

Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768

Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103

Schmidt K. B., et al., 2014, ApJ, 786, 57

Schramm D. N., Turner M. S., 1998, Reviews of Modern Physics, 70, 303

Scott N., et al., 2018, MNRAS, 481, 2299

Searle L., 1971, ApJ, 168, 327

Seaton M. J., 1960, Reports on Progress in Physics, 23, 313

Shapley H., 1918, PASP, 30, 42

Shapley A. E., et al., 2019, arXiv e-prints, p. arXiv:1907.07189

Sharp R. G., Bland-Hawthorn J., 2010, ApJ, 711, 818

Sharp R., et al., 2015, MNRAS, 446, 1551

Shaw R. A., Dufour R. J., 1995, PASP, 107, 896

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Simons R. C., et al., 2020, arXiv e-prints, p. arXiv:2011.03553

Skibba R. A., et al., 2014, ApJ, 784, 128

Skillman E. D., Kennicutt R. C., Hodge P. W., 1989, ApJ, 347, 875

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15

Springel V., et al., 2005, Nature, 435, 629

Stanway E. R., Eldridge J. J., 2018, MNRAS, 479, 75

Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJL, 462, L17

Steidel C. C., et al., 2014, ApJ, 795, 165

Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, ApJ, 826, 159

Stevance H., Eldridge J., Stanway E., 2020, The Journal of Open Source Software, 5, 1987

Storey P. J., Sochi T., Badnell N. R., 2014, MNRAS, 441, 3028

Stott J. P., et al., 2013, MNRAS, 436, 1130

Stott J. P., et al., 2016, MNRAS, 457, 1888

Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, ApJ, 836, 164

Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., 2018, ApJ, 868, 117

Strömgren B., 1939, ApJ, 89, 526

Sutherland R. S., Dopita M. A., 2017, ApJS, 229, 34

Taylor E. N., et al., 2011, MNRAS, 418, 1587

Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1

The Astropy Collaboration et al., 2018, AJ, 156, 123

Thomas A. D., Dopita M. A., Kewley L. J., Groves B. A., Sutherland R. S., Hopkins A. M., Blanc G. A., 2018, ApJ, 856, 89

Timmes F. X., Woosley S. E., Weaver T. A., 1995, ApJS, 98, 617

Tissera P. B., Rosas-Guevara Y., Bower R. G., Crain R. A., del P Lagos C., Schaller M., Schaye J., Theuns T., 2019, MNRAS, 482, 2208

Topping M. W., Shapley A. E., Reddy N. A., Sanders R. L., Coil A. L., Kriek M., Mobasher B., Siana B., 2020a, MNRAS, 495, 4430

Topping M. W., Shapley A. E., Reddy N. A., Sanders R. L., Coil A. L., Kriek M., Mobasher B., Siana B., 2020b, MNRAS, 499, 1652

Torrey P., Cox T. J., Kewley L., Hernquist L., 2012, ApJ, 746, 108

Tremonti C. A., et al., 2004, ApJ, 613, 898

Trenti M., Stiavelli M., 2008, ApJ, 676, 767

Trenti M., et al., 2011, ApJL, 727, L39

Trenti M., et al., 2012, ApJ, 746, 55

Treu T., et al., 2015, ApJ, 812, 114

Troncoso P., et al., 2014, A&A, 563, A58

Tumlinson J., et al., 2011, Science, 334, 948

Tyson J. A., Valdes F., Wenk R. A., 1990, ApJL, 349, L1

Vale Asari N., Stasińska G., Morisset C., Cid Fernandes R., 2016, MNRAS, 460, 1739

Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, MNRAS, 404, 1639

Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295

Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A, 43, 769

Viironen K., et al., 2018, A&A, 614, A129

Vila-Costas M. B., Edmunds M. G., 1992, MNRAS, 259, 121

Virtanen P., et al., 2020, Nature Methods, 17, 261

Wake D. A., et al., 2011, ApJ, 728, 46

Wanajo S., Sekiguchi Y., Nishimura N., Kiuchi K., Kyutoku K., Shibata M., 2014, ApJL, 789, L39

Wang Y., Yang X., Mo H. J., van den Bosch F. C., 2007, ApJ, 664, 608

Wang X., et al., 2019, ApJ, 882, 94

Wang X., et al., 2020, ApJ, 900, 183

Weiner B. J., et al., 2009, ApJ, 692, 187

Werk J. K., Prochaska J. X., Thom C., Tumlinson J., Tripp T. M., O'Meara J. M., Peeples M. S., 2013, ApJS, 204, 17

Westmoquette M. S., Smith L. J., Gallagher J. S. I., 2011, MNRAS, 414, 3719

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wuyts E., Rigby J. R., Sharon K., Gladders M. D., 2012, ApJ, 755, 73

Wuyts E., et al., 2014, ApJL, 789, L40

Wuyts E., et al., 2016, ApJ, 827, 74

Yabe K., et al., 2012, PASJ, 64, 60

Yabe K., et al., 2014, MNRAS, 437, 3647

Yates R. M., Kauffmann G., Guo Q., 2012, MNRAS, 422, 215

Yates R. M., Schady P., Chen T. W., Schweyer T., Wiseman P., 2020, A&A, 634, A107

York D. G., et al., 2000, AJ, 120, 1579

Yuan T. T., Kewley L. J., Swinbank A. M., Richard J., Livermore R. C., 2011, ApJL, 732, L14

Yuan T.-T., Kewley L. J., Richard J., 2013, ApJ, 763, 9

Yuan T., et al., 2020, Nature Astronomy,

Zahid H. J., Kewley L. J., Bresolin F., 2011, ApJ, 730, 137

Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, ApJL, 771, L19

Zahid H. J., Dima G. I., Kudritzki R.-P., Kewley L. J., Geller M. J., Hwang H. S., Silverman J. D., Kashino D., 2014a, ApJ, 791, 130

Zahid H. J., et al., 2014b, ApJ, 792, 75

Zaritsky D., Kennicutt Jr. R. C., Huchra J. P., 1994, ApJ, 420, 87

Zehavi I., et al., 2005, ApJ, 630, 1

Zehavi I., et al., 2011, ApJ, 736, 59

Zhang K., et al., 2017, MNRAS, 466, 3217

Zhou L., et al., 2017, MNRAS, 470, 4573

de la Torre S., et al., 2011, MNRAS, 412, 825

pandas development team T., 2020, pandas-dev/pandas: Pandas, doi:10.5281/zenodo.3509134, `https://doi.org/10.5281/zenodo.3509134`

van der Velden E., 2020, The Journal of Open Source Software, 5, 2004